

Pathwise Stieltjes integrals of discontinuously evaluated stochastic processes*

Zhe Chen[†] Lasse Leskelä[‡] Lauri Viitasaari[§]

December 5, 2016

Abstract

In this article we study the existence of pathwise Stieltjes integrals of the form $\int f(X_t) dY_t$ for nonrandom, possibly discontinuous, evaluation functions f and Hölder continuous random processes X and Y . We discuss a notion of sufficient variability for the process X which ensures that the paths of the composite process $t \mapsto f(X_t)$ are almost surely regular enough to be integrable. We show that the pathwise integral can be defined as a limit of Riemann–Stieltjes sums for a large class of discontinuous evaluation functions of locally finite variation, and provide new estimates on the accuracy of numerical approximations of such integrals, together with a change of variables formula for integrals of the form $\int f(X_t) dX_t$.

Keywords: composite stochastic process, generalized Stieltjes integral, fractional calculus, Riemann–Liouville integral, fractional Sobolev space, Gagliardo–Slobodeckij seminorm, fractional Sobolev–Slobodeckij space, bounded p -variation

MSC 2010: 60H05, 60G22, 26A33, 60G07, 60G15

1 Introduction

In this article we study the existence of pathwise Stieltjes integrals of the form

$$\int_0^T f(X_t) dY_t, \quad (1.1)$$

where X and Y are real-valued stochastic processes with time parameter $t \in [0, T]$, and the values of X are evaluated using a nonrandom function $f : \mathbb{R} \rightarrow \mathbb{R}$. If X is adapted and Y is a semimartingale with respect to some common filtration, then the integral can be handled using classical Itô calculus

*A substantial part of this research has been supported by the Emil Aaltonen Foundation.

[†]New York University, Tandon School of Engineering, Dept. Finance and Risk Engineering

[‡]Aalto University, School of Science, Dept. Mathematics and Systems Analysis

[§]Aalto University, School of Electrical Engineering, Dept. Communications and Networking

techniques. On the other hand, if the function f and the paths of X and Y are Hölder continuous of sufficiently large order, then the above integral exists almost surely as a standard Riemann–Stieltjes integral, as a consequence of the classical results of L. C. Young [You36, DN98]. However, in several application contexts it is important to consider stochastic processes which are not semimartingales and evaluation functions which are not continuous. For example, in bang-bang stochastic control [KD92] and stop-loss trading [FS04, Son06] the function f corresponds to the indicator function of a certain threshold. In this case the paths of $t \mapsto f(X_t)$ are typically discontinuous, and even worse, of infinite p -variation for all $p \geq 1$, which rules out rough path techniques introduced in [Lyo98, FH14].

Motivated by stochastic models where X and Y are non-semimartingales and f is discontinuous, we investigate conditions which allow to define the integral (1.1) pathwise. The key idea in our approach is the assumption that *sufficient variability* in the probability distribution of X makes it unlikely for X to spend much time in the discontinuity points of f and renders the composite process $t \mapsto f(X_t)$ regular enough to be integrable. Our main contribution is to quantify the notion of sufficient variability and show that the integral (1.1) can be defined as a limit of Riemann–Stieltjes sums for a large class of discontinuous evaluation functions of locally finite variation. We also provide new estimates on the accuracy of numerical approximations of such integrals computed using Riemann–Stieltjes sums over finite dense partitions, together with a change of variables formula for integrals of the form $\int f(X_t) dX_t$.

The main technical tool in our analysis is the pathwise fractional Stieltjes integral introduced by Zähle in [Zäh98], which is based on reinterpreting the Stieltjes integral as the Lebesgue integral of a product of fractional derivatives. The main technical part of the article is to show that under sufficient regularity conditions, the paths of $f(X_t)$ have almost surely finite Gagliardo seminorms and fit into fractional Sobolev spaces where fractional derivatives behave well [DNPV12]. Consequently, the integral (1.1) exists as a fractional Stieltjes integral. With this result at hand, we show how ordinary Riemann–Stieltjes sums approximating (1.1) can also be interpreted as fractional Stieltjes integrals, and the convergence of Riemann–Stieltjes sums is then obtained by showing that suitably interpolated paths of $f(X_t)$ converge to the actual paths in the corresponding fractional Sobolev space.

Since the seminal paper of Zähle [Zäh98], fractional Stieltjes integrals of stochastic processes have been actively studied, especially for fractional Brownian motions. Nualart and Răşcanu [NR02] used fractional calculus techniques to study differential equations driven by Hölder continuous processes; all integrals of the form (1.1) were restricted to Lipschitz continuous evaluation functions f . To the best of our knowledge, the first paper to discuss integrals of the form (1.1) for discontinuous evaluation functions is [AMV10], where similar types of questions as in this paper were studied for geometric fractional Brownian motions. Unfortunately, some proofs of the key results in [AMV10] contain flaws that appear difficult to fix without incorporating additional assumptions.

Tikanmäki [Tik12] extended the analysis to nonlinear integral functionals of fractional Brownian motions. Another closely related paper is [AV15], where the rate of convergence of forward Riemann–Stieltjes sums over uniform partitions were studied for fractional Brownian motions. All of the aforementioned papers and most others focus on fractional Brownian motions of Hurst index $H > 1/2$. However, see [BN11] for a special case with $1/3 < H < 1/2$.

In contrast with the aforementioned works, our current paper studies the integrability questions in a general framework with minimal assumptions, not restricted to fractional Brownian motions nor even Gaussian processes. Besides being more general, the key results, when restricted to fractional Brownian motions, turn out also to be stronger (see Example 3.11).

The rest of the paper is organized as follows. Section 2 introduces the key assumptions, Section 3 contains the main results, and Section 4 contains the proofs. Some of the more technical proofs are deferred into separate appendices in the end of the paper.

2 Key assumptions

2.1 Sufficient variability

We make the following assumption for a random process $X : [0, T] \rightarrow \mathbb{R}$ which guarantees that the probability distribution of X_t is not too much concentrated around any single value.

Assumption 2.1. For almost every $t \in [0, T]$, the random variable X_t admits a probability density function $y \mapsto p_t(y)$ that is bounded according to

$$\sup_y p_t(y) \leq \hat{p}_t, \quad (2.1)$$

where $\int_0^T \hat{p}_t dt < \infty$.

This assumption holds for a broad variety of random processes, for example:

- If X is a stationary process such that X_0 has a bounded probability density function, then (2.1) holds with \hat{p}_t being a constant.
- If X is a Gaussian process with a covariance function $R(s, t) = \text{Cov}(X_s, X_t)$ such that

$$\int_0^T R(t, t)^{-1/2} dt < \infty,$$

then (2.1) holds with $\hat{p}_t = R(t, t)^{-1/2}$. Especially, Brownian motions and all fractional Brownian motions with Hurst parameter $H \in (0, 1)$ satisfy this.

Moreover, the class of random processes satisfying Assumption 2.1 is closed with respect to sums of stochastically independent processes and shifts by nonrandom functions. More precisely:

- If X satisfies Assumption 2.1, then so does $t \mapsto X_t + m_t$, for any measurable nonrandom function $t \mapsto m_t$.
- If $X^{(1)}$ and $X^{(2)}$ are independent random processes such that $X_t^{(i)}$ admits a probability density function $p_t^{(i)}$ for almost every $t \in [0, T]$, and if $X^{(1)}$ satisfies Assumption 2.1, then so does $X_t = X_t^{(1)} + X_t^{(2)}$, with probability density function bounded by

$$p_t = \int_{\mathbb{R}} p_t^{(1)}(z - y) p_t^{(2)}(y) dy \leq \hat{p}_t^{(1)} \int_{\mathbb{R}} p_t^{(2)}(y) dy = \hat{p}_t^{(1)}.$$

2.2 Continuity

Most of the underlying random process studied in this paper will be assumed to have almost surely Hölder continuous paths. In addition, we shall often require the following standard continuity assumption on the time differences of X which is the classical Kolmogorov–Chentsov criterion [Kal02, Theorem 3.23] for the Hölder continuity of X of order $\alpha < (\lambda - 1)/\delta$.

Assumption 2.2. There exist numbers $\delta > 1$ and $\lambda > 1$ such that

$$\sup_{0 \leq s < t \leq T} \frac{E|X_t - X_s|^\delta}{|t - s|^\lambda} < \infty.$$

Remark 2.3. Any Gaussian process with Hölder continuous paths of order $\alpha > 0$ satisfies Assumption 2.2 for any $\delta > 1$ and any $\lambda < \alpha\delta$, see [ASVY14].

3 Results

3.1 Pathwise integrability

Throughout this section we denote by $(\text{ZS}) \int$ the fractional Zähl–Stieltjes integral which we formally recall in Section 4.1.

Theorem 3.1. *Let X and Y be Hölder continuous random processes of orders α and β , respectively, such that $\alpha + \beta > 1$. Assume also that X admits a probability density function bounded according to Assumption 2.1. Then for any $f : \mathbb{R} \rightarrow \mathbb{R}$ of locally finite variation such that $f(X_{0+})$ exists almost surely, the pathwise integral*

$$(\text{ZS}) \int_0^T f(X_t) dY_t$$

exists almost surely.

Remark 3.2. If the random variable X_t admits a probability density function at every $t \in [0, T]$, then the assumption that $f(X_{0+})$ exists almost surely is automatically valid, because the set of discontinuities of f is countable.

Remark 3.3. By examining the proof of Theorem 3.1 (actually its key argument in the proof of Proposition 4.3) it is evident that a sufficient condition for the existence of the integral is that

$$\sup_{y \in \mathbb{R}} E \int_0^T |X_t - y|^{-\theta/\alpha} dt < \infty \quad (3.1)$$

for some $\theta \in (1 - \beta, \alpha)$. This condition means that while X may (and does in our cases of interest) cross some levels infinitely often, the crossings do not occur too often in the above sense. Furthermore, Assumption 2.1 guarantees that (3.1) holds for process X . Consequently, the integral exists for all Hölder continuous nonrandom processes $x : [0, T] \rightarrow \mathbb{R}$ satisfying

$$\sup_{y \in \mathbb{R}} \int_0^T |x_t - y|^{-\theta/\alpha} dt < \infty.$$

3.2 Change of variables

Theorem 3.4. *Let X be a Hölder continuous random process of order $\alpha > 1/2$, with a probability density function bounded according to Assumption 2.1. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be absolutely continuous, having a derivative f' of locally finite variation such that $f'(X_{0+})$ exists almost surely. Then for any $t \in [0, T]$,*

$$f(X_t) - f(X_0) = (\text{ZS}) \int_0^t f'(X_s) dX_s \quad (3.2)$$

almost surely.

3.3 Riemann–Stieltjes integrability

A tagged partition of $[0, T]$ is a sequence $\pi = (t_0, \dots, t_k; \xi_1, \dots, \xi_k)$ such that $0 = t_0 < t_1 < \dots < t_k$ and $\xi_i \in [t_{i-1}, t_i]$ for all i . The mesh of π is denoted by $\|\pi\| = \max_i (t_i - t_{i-1})$, and the Riemann–Stieltjes sum of f against g along π by

$$\text{RS}(f, g, \pi) = \sum_{i=1}^k f(\xi_i)(g(t_i) - g(t_{i-1})).$$

If $\text{RS}(f, g, \pi_n)$ converges to a unique limit for any sequence of tagged partitions such that $\|\pi_n\| \rightarrow 0$, then this limit is denoted $(\text{RS}) \int_0^T f_t dg_t$ and called the Riemann–Stieltjes integral of f against g .

The next result concerns the Riemann–Stieltjes integrability of random processes and can be viewed as the main result of this paper. We denote convergence in probability by \xrightarrow{P} .

Theorem 3.5. *Let X and Y be Hölder continuous of orders α and β , respectively, such that $\alpha + \beta > 1$. Assume also that X_t admits a probability density function p_t at every $t \in (0, T)$ bounded according to Assumption 2.1 and*

$$\sup_{0 \leq s < t \leq T} \frac{E|X_t - X_s|^\delta}{|t - s|^\lambda} < \infty$$

for some $\delta > 1$ and $\lambda > 1$ such that $\frac{\lambda}{1+\delta} + \beta > 1$. In addition, assume that there exists a finite constant c_0 such that either:

- (i) $p_t(x) \leq c_0$ for all $x \in \mathbb{R}$ and all $t \in (0, T)$, and $f(X_{0+})$ exists almost surely; or
- (ii) $p_t(x) \leq c_0$ for all $|x| \geq \epsilon/2$ and all $t \in (0, T)$, and f restricted to $(-\epsilon, \epsilon)$ is Lipschitz continuous for some $\epsilon > 0$.

Then there exists a random variable $(RS) \int_0^T f(X_t) dY_t$ such that

$$RS(f \circ X, g, \pi_n) \xrightarrow{P} (RS) \int_0^T f(X_t) dY_t \quad (3.3)$$

for any sequence of tagged partitions with $\|\pi_n\| \rightarrow 0$, and

$$(RS) \int_0^T f(X_t) dY_t = (ZS) \int_0^T f(X_t) dY_t \quad a.s.$$

Furthermore, the convergence in (3.3) holds almost surely for any sequence of tagged partitions satisfying

$$\sum_{n=1}^{\infty} \|\pi_n\|^a < \infty$$

for some $a < \frac{\lambda}{1+\delta} + \beta - 1$.

Remark 3.6. Item (i) in the assumption corresponds to a simpler case where the density is uniformly bounded also in t , and in this case we do not require additional regularity on f around X_0 . For example, this is the case if X is a stationary process with bounded density.

Remark 3.7. Item (ii) in the assumption corresponds to a case of random process X such that $X_0 = 0$ in which case the assumption is usually satisfied as the density p_t of the process X_t approaches to a Dirac mass at zero when t approaches to zero. However, as a price the continuity of f at $X_0 = 0$ is not sufficient but we need additional regularity. For example, assumption on the density in item (ii) is valid for all Gaussian process X such that $X_0 = 0$ and $R(t, t) > 0$ for all $t \in (0, T)$. In a general case where X_0 is some arbitrary constant, then the assumption can be replaced with Lipschitz continuity of f in a small neighbourhood of X_0 and the density should be bounded for all $|x - X_0| \geq \epsilon/2$.

3.3.1 Dyadic partitions

The following result considers the rate of convergence for the almost sure convergence in the case of dyadic partitions.

Theorem 3.8. *Let π_n be a sequence of tagged dyadic partitions, i.e. $\pi_n = \{\frac{kT}{2^n}, k = 0, \dots, 2^n\}$. Suppose further that the assumptions of Theorem 3.5 prevail. Then there exists an almost surely finite random variable $C(\omega)$ such that, for any $\epsilon > 0$, we have*

$$\left| \text{RS}(f \circ X, g, \pi_n) - (\text{RS}) \int_0^T f(X_t) dY_t \right| \leq C \left(\frac{1}{2} \right)^{\left[\frac{\lambda}{1+\delta} - \theta - \epsilon \right] n}.$$

3.3.2 Accuracy of interpolations

For the following result, we need to rule out functions of locally finite variation for which the variation $V_{[-k,k]}(f)$ over the interval $[-k, k]$ grows too fast as $k \rightarrow \infty$. A sufficient condition is to assume that the growth rate of $V_{[-k,k]}(f)$ is slow compared to the decay rate of the tail probability distribution of the maximum of X over $[0, T]$, in the sense that

$$\sum_{k=1}^{\infty} V_{[-k,k]}(f) P(\|X\|_{\infty} > k-1)^{\frac{1}{p}} < \infty \quad (3.4)$$

for all $p \geq 1$.

Theorem 3.9. *Suppose that assumptions of Theorem 3.5 prevail with $\alpha, \beta \in (0, 1)$ and $\delta, \lambda > 1$ such that $\alpha + \beta > 1$ and $\frac{\lambda}{1+\delta} + \beta > 1$. Assume that in addition, the Hölder constant of Y is bounded by $E[Y]_{\beta, \infty}^p < \infty$ for all $p \geq 1$. Assume further that either f has finite variation, or f satisfies (3.4). Then for any $a < \frac{\lambda}{1+\delta} + \beta - 1$ there exists a finite constant c such that*

$$E \left| \text{RS}(f \circ X, g, \pi) - (\text{RS}) \int_0^T f(X_t) dY_t \right| \leq c \|\pi\|^a \quad (3.5)$$

for all tagged partitions of $[0, T]$ with mesh $\|\pi\| \leq 1$.

Remark 3.10. If Y is a Gaussian process on $[0, T]$ with Hölder continuous paths of order β , then the Hölder constant of Y is bounded by $E[Y]_{\beta, \infty}^p < \infty$ for all $p \geq 1$, as required above [ASVY14].

Example 3.11. Let X and Y be (possibly dependent) fractional Brownian motions with Hurst indices H_1 and H_2 , respectively, such that $H_1 + H_2 > 1$, and let f be of locally finite variation and Lipschitz continuous in a neighborhood $(-\epsilon, \epsilon)$ of zero. Then the probability density function of X_t is bounded by

$$p_t(x) \leq (2\pi)^{-1/2} t^{-H_1} \exp \left(-\frac{(\epsilon/2)^2}{2t^{2H_1}} \right) \quad \text{for } |x| \geq \epsilon/2,$$

and the right side above is bounded over $t \in (0, T)$. This shows that the condition (ii) in Theorem 3.5 is valid. Moreover, Assumption 2.1 is valid for X because $p_t \leq (2\pi)^{-1/2} t^{-H_1}$ for all $t \in (0, T)$.

Given an arbitrary $a < H_1 + H_2 - 1$, choose $\alpha < H_1$ and $\beta < H_2$ so that $\alpha + \beta - 1 > a$. then X and Y have Hölder continuous paths of orders α and

β , respectively. Furthermore, by Remark 2.3, we may choose $\lambda, \delta > 1$ so that X satisfies Assumption 2.2 and $\frac{\lambda}{1+\delta}$ is so close to α that $a < \frac{\lambda}{1+\delta} + \beta - 1$. Furthermore, the supremum of a Gaussian process has Gaussian tails [Sam91], and consequently (3.4) is satisfied for all f for which $V_{[-k,k]}(f)$ has at most exponential growth. Finally, because Y is a Gaussian process with Hölder continuous paths of order β , then $E[Y]_{\beta,\infty}^p < \infty$ for all $p \geq 1$ [ASVY14].

Hence the triplet (X, Y, f) fulfills all assumptions required in Theorems 3.5 and 3.9, and we may conclude that

$$E \left| \text{RS}(f \circ X, g, \pi) - (\text{RS}) \int_0^T f(X_t) dY_t \right| \leq c \|\pi\|^{H_1+H_2-1-\epsilon}$$

for all $\epsilon = H_1 + H_2 - 1 - a > 0$. A similar type of result for $H_1 = H_2 > 1/2$, restricted to forward Riemann–Stieltjes sums over a uniform mesh, was discovered in [AV15, Theorem 3.1], but with a significantly worse rate of $\|\pi\|^{H-1/2-\epsilon}$.

Example 3.12. Let X be a stationary α -Hölder continuous Gaussian processes, e.g. a fractional Ornstein–Uhlenbeck process as in [CKM03], and Y be a β -Hölder continuous Gaussian process such that $\alpha + \beta > 1$. Let f be an arbitrary function of locally finite variation such that the variation of f over large intervals grows at most exponentially. Then the triplet (f, X, Y) satisfies the assumptions of Theorems 3.5 and 3.9, with (i) in place of (ii). Then

$$E \left| \text{RS}(f \circ X, g, \pi) - (\text{RS}) \int_0^T f(X_t) dY_t \right| \leq c \|\pi\|^{\alpha+\beta-1-\epsilon}$$

for all $\epsilon > 0$.

4 Proofs

4.1 Fractional integrals and derivatives

4.1.1 Riemann–Liouville operators

Let us recall some basic facts and notations of fractional integrals and derivatives from [SKM93]. Let $(a, b) \subset \mathbb{R}$ be a nonempty bounded interval. For $p \in [1, \infty)$ (resp. $p = \infty$), we denote by $L_p = L_p(a, b)$ the space of p -integrable (resp. essentially bounded) real functions on (a, b) . The fractional left and right Riemann–Liouville integrals of order $\alpha > 0$ of a function $f \in L_1$ are denoted¹ by

$$I_{a+}^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_a^t \frac{f(s)}{(t-s)^{1-\alpha}} ds$$

and

$$I_{b-}^\alpha f(t) = \frac{(-1)^{-\alpha}}{\Gamma(\alpha)} \int_t^b \frac{f(s)}{(t-s)^{1-\alpha}} ds.$$

¹Here we follow Zähle’s (and Liouville’s) convention of including $(-1)^{-\alpha} = e^{-i\pi\alpha}$ into the definition of right-sided fractional integrals [Zäh98].

The above integrals converge for almost every $t \in (a, b)$, and $I_{a+}^\alpha f$ and I_{b-}^α may be considered as functions in L_1 . By convention, I_{a+}^0 and I_{b-}^0 are defined as the identity operator. Moreover, the integral operators $I_{a+}^\alpha, I_{b-}^\alpha : L_1 \rightarrow L_1$ are linear and one-to-one. The inverse operators are known as Riemann–Liouville fractional derivatives, and denoted by $I_{a+}^{-\alpha} = (I_{a+}^\alpha)^{-1}$ and $I_{b-}^{-\alpha} = (I_{b-}^\alpha)^{-1}$. Because $L_p(a, b) \subset L_1(a, b)$ for every $p \in [1, \infty]$, we may also consider the fractional integrals as linear operators from L_p into L_1 , with ranges denoted by $I_{a+}^\alpha(L_p)$ and $I_{b-}^\alpha(L_p)$.

For any $\alpha \in (0, 1)$ and for any $f \in I_{a+}^\alpha(L_1)$ and $g \in I_{b-}^\alpha(L_1)$, the Weyl–Marchaud derivatives

$$D_{a+}^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(t)}{(t-a)^\alpha} + \alpha \int_a^t \frac{f(t) - f(s)}{(t-s)^{\alpha+1}} ds \right)$$

and

$$D_{b-}^\alpha g(t) = \frac{(-1)^\alpha}{\Gamma(1-\alpha)} \left(\frac{g(t)}{(b-t)^\alpha} + \alpha \int_t^b \frac{g(t) - g(s)}{(s-t)^{\alpha+1}} ds \right)$$

are well defined, and coincide with the Riemann–Liouville derivatives according to $D_{a+}^\alpha f(t) = I_{a+}^{-\alpha} f(t)$ and $D_{b-}^\alpha g(t) = I_{b-}^{-\alpha} g(t)$ for almost every $t \in (a, b)$.

4.1.2 Fractional Zähle–Stieltjes integrals

Let f and g be measurable real functions defined on some set containing the interval (a, b) . Assume that the limits $f(a+), g(a+), g(b-)$ exist in \mathbb{R} , and denote $f_{a+}(t) = f(t) - f(a+)$ and $g_{b-}(t) = g(t) - g(b-)$. When $f_{a+} \in I_{a+}^\alpha(L_p)$ and $g_{b-} \in I_{b-}^{1-\alpha}(L_q)$ for some $\alpha \in [0, 1]$ and $p, q \in [1, \infty]$ such that $1/p + 1/q = 1$, the fractional version of the Stieltjes integral introduced by Zähle [Zäh98] is defined by

$$\begin{aligned} (\text{ZS}) \int_a^b f_t dg_t &= (-1)^\alpha \int_a^b D_{a+}^\alpha (f - f(a+))(t) D_{b-}^{1-\alpha} (g - g(b-))(t) dt \\ &\quad + f(a+)(g(b-) - g(a+)), \end{aligned} \quad (4.1)$$

where the right side does not depend on α .

4.1.3 Fractional Sobolev–Slobodeckij spaces

To help treating fractional integrals and derivatives in a convenient way, we introduce the following fractional Sobolev–Slobodeckij type space below. For $\alpha > 0$, we denote by $W_1^\alpha(a+)$ the space of measurable functions $f : (a, b) \rightarrow \mathbb{R}$ such that $f(a+) \in \mathbb{R}$ exists and

$$\|f\|_{W_1^\alpha(a+)} = \int_a^b \frac{|f(t)|}{(t-a)^\alpha} dt + \int_a^b \int_a^t \frac{|f(t) - f(s)|}{|t-s|^{1+\alpha}} ds dt$$

is finite. We denote by $W_\infty^\alpha(b-)$ the space of measurable functions $f : (a, b) \rightarrow \mathbb{R}$ such that $f(b-) \in \mathbb{R}$ exists and

$$\|f\|_{W_\infty^\alpha(b-)} = \sup_{t \in (a, b)} \frac{|f(b-) - f(t)|}{(b-t)^\alpha} + \sup_{t \in (a, b)} \int_t^b \frac{|f(t) - f(s)|}{|t-s|^{1+\alpha}} ds$$

is finite. The spaces $W_1^\alpha(a+)$ and $W_\infty^\alpha(b-)$ contain all Hölder continuous functions of order $\alpha' > \alpha$. More importantly, these spaces also contain discontinuous functions of infinite power variation, as we will see later.

The following result may be considered as a mild generalization of a similar estimate in [NR02], where a subspace of Hölder continuous functions was employed in place of $W_\infty^\alpha(b-)$. See also [Zäh01, Theorem 1.1] for a similar result for square integrable functions.

Proposition 4.1. *Assume that $f \in W_1^\alpha(a+)$ and $g \in W_\infty^{1-\alpha}(b-)$ for some $\alpha \in (0, 1)$. Then the integral in (4.1) is well defined, representable as*

$$(ZS) \int_a^b f_t dg_t = (-1)^\alpha \int_a^b D_{a+}^\alpha f(t) D_{b-}^{1-\alpha}(g - g(b-))(t) dt, \quad (4.2)$$

and bounded by

$$\left| (ZS) \int_a^b f_t dg_t \right| \leq \frac{\|f\|_{W_1^\alpha(a+)} \|g\|_{W_\infty^{1-\alpha}(b-)}}{\Gamma(\alpha)\Gamma(1-\alpha)}. \quad (4.3)$$

Proof. The definitions of $W_1^\alpha(a+)$ and $W_\infty^{1-\alpha}(b-)$ guarantee that the Weyl–Marchaud derivatives $D_{a+}^\alpha f$ and $D_{b-}^{1-\alpha}(g - g(b-))$ are well defined and satisfy

$$\|D_{a+}^\alpha f\|_{L_1} \leq \frac{\|f\|_{W_1^\alpha(a+)}}{\Gamma(1-\alpha)}$$

and

$$\|D_{b-}^{1-\alpha}(g - g(b-))\|_{L_\infty} \leq \frac{\|g\|_{W_\infty^{1-\alpha}(b-)}}{\Gamma(\alpha)}.$$

With the help of [SKM93, Theorem 13.2], one can verify that in this case $f \in I_{a+}^\alpha(L_1)$ and $g - g(b-) \in I_{b-}^{1-\alpha}(L_\infty)$. Then (4.2) follows by [Zäh98, Eq. (22')], and (4.3) follows by plugging in the above two estimates into (4.2). \square

4.2 Hölder and Gagliardo seminorms

The Hölder seminorm of order $\alpha > 0$ of a measurable function $x : [0, T] \rightarrow \mathbb{R}$ is denoted by

$$[x]_{\alpha, \infty} = \sup_{0 \leq s < t \leq T} \frac{|x(t) - x(s)|}{|t - s|^\alpha},$$

and the Gagliardo seminorm of order $\alpha > 0$ and exponent $p \in [1, \infty)$ by

$$[x]_{\alpha, p} = \left(\int_0^T \int_0^T \frac{|x(t) - x(s)|^p}{|t - s|^{1+\alpha p}} ds dt \right)^{1/p}. \quad (4.4)$$

Lemma 4.2. *For any $\alpha > 0$ and $\epsilon > 0$,*

$$\|x\|_{W_1^\alpha(0+)} \leq (1 - \alpha)^{-1} T^{1-\alpha} \|x\|_\infty + \frac{1}{2} [x]_{\alpha, 1},$$

$$\|x\|_{W_\infty^\alpha(T-)} \leq (1 + \epsilon^{-1}) T^\epsilon [x]_{\alpha+\epsilon, \infty},$$

and

$$[x]_{\alpha, 1} \leq \epsilon^{-1} (1 + \epsilon)^{-1} T^{1+\epsilon} [x]_{\alpha+\epsilon, \infty}.$$

Proof. These are straightforward computations. \square

4.3 Regularity of composite paths

In this section we analyze the regularity of $t \mapsto f \circ X(t)$ when f is a function of locally finite variation and $t \mapsto X(t)$ is a random or nonrandom Hölder continuous function. The main result is the following.

Proposition 4.3. *Let X be a Hölder continuous random process of order $\alpha > 0$, and such that Assumption 2.1 holds. Then for any function f of locally finite variation, $[f \circ X]_{\theta,p} < \infty$ almost surely for all $\theta \in (0, 1)$ and $p \in [1, \infty)$ such that $1 \leq p < \alpha/\theta$.*

Any function $f : \mathbb{R} \rightarrow \mathbb{R}$ of locally finite variation can be written as a sum of a right-continuous and a left-continuous function of locally finite variation [Kal02, Prop. 2.19]. This is why in the sequel we will mainly discuss right-continuous functions. If we ignore the value of f at its jump points, then we may define the right-continuous function $f_r(x) = \lim_{h \downarrow 0} f(x+h)$, and take this function as a representative of f .

Lemma 4.4. *Assume that $\text{Leb}\{t \in [0, T] : x_t \in J_f\} = 0$ where J_f denotes the set of discontinuities of f . Then for any $\theta \in (0, 1)$ and $p \in [1, \infty)$,*

$$[f \circ x]_{\theta,p} = [f_r \circ x]_{\theta,p}.$$

Proof. The claim follows by noting that

$$\frac{|f(X_t) - f(X_s)|^p}{|t - s|^{1+\theta p}} = \frac{|f_r(X_t) - f_r(X_s)|^p}{|t - s|^{1+\theta p}}$$

for all (s, t) outside the set $\{(s, t) : x_s \in J_f \text{ or } x_t \in J_f\}$ which has Leb^2 -measure zero by assumption. \square

Note that because J_f is countable for any f of locally finite variation, another sufficient condition for the above result is to assume that $\text{Leb}\{t : x_t = y\} = 0$ for all y .

Lemma 4.5. *If $\theta > 0$ and $x : [0, T] \rightarrow \mathbb{R}$ is Hölder continuous of order $\alpha > 0$, then*

$$\int_0^t \frac{1_{\{x_s < y\}}}{(t-s)^{1+\theta}} ds \leq \theta^{-1} [x]_{\alpha,\infty}^{\theta/\alpha} |x_t - y|^{-\theta/\alpha}$$

for all $t \in [0, T]$ and for all $y < x_t$.

Proof. Fix $t \in [0, T]$ and $y < x_t$, and denote by I the integral above. If $x_s \geq y$ for all $s \in [0, t]$, then $I = 0$ and there is nothing to prove. Let us consider the case where $x_s < y$ for at least one $s \in [0, t]$. Let $s^* = \sup\{s \in [0, t] : X_s < y\}$ be the supremum of the time indices for which this happens. Then $s^* \in [0, t]$ and $x_{s^*} = y$ by the continuity of x . In this case the integral is bounded from above by

$$I \leq \int_0^{s^*} \frac{1}{(t-s)^{1+\theta}} ds = \frac{(t-s^*)^{-\theta} - t^{-\theta}}{\theta} \leq \theta^{-1} (t-s^*)^{-\theta}. \quad (4.5)$$

By the Hölder continuity of x , it follows that

$$|x_t - y| = |x_t - x_{s^*}| \leq [x]_{\alpha, \infty} |t - s^*|^\alpha,$$

so that

$$I \leq \theta^{-1} (t - s^*)^{-\theta} \leq \theta^{-1} [x]_{\alpha, \infty}^{\theta/\alpha} |x_t - y|^{-\theta/\alpha}.$$

□

For a right-continuous function f of locally finite variation, denote its variation measure by μ_f . This is the unique locally finite Borel measure on \mathbb{R} such that the variation of f over the interval $[a, b]$ equals $V_{(a, b)}(f) = \mu_f(a, b]$ for all $a < b$.

Proposition 4.6. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be right-continuous and of finite variation. Let $x : [0, T] \rightarrow \mathbb{R}$ be Hölder continuous of order $\alpha > 0$. Then for any $\theta \in (0, 1)$ and any $p \in [1, \infty)$,*

$$[f \circ x]_{\theta, p}^p \leq 2^p (\theta p)^{-1} \mu_f(K_x)^{p-1} [x]_{\alpha, \infty}^{\theta p/\alpha} \int_0^T \int_{K_x} |x_t - y|^{-\theta p/\alpha} \mu_f(dy) dt, \quad (4.6)$$

where K_x is the closure of the range of x .

Proof. (i) Let us first assume that f is increasing and right-continuous. In this case the variation measure μ_f equals the Lebesgue–Stieltjes measure of f and satisfies $\mu(x_1, x_2] = f(x_2) - f(x_1)$ for $x_1 < x_2$. Let $G_x = \{(t, x_t) : t \in [0, T]\}$ be the graph of x . We may and will assume that $(\text{Leb} \otimes \mu_f)(G_x) = 0$ because otherwise the right side of (4.6) is infinite and there is nothing to prove.

Observe next that for all real numbers x_1 and x_2 in $K = K_x$,

$$f(x_2) - f(x_1) = \begin{cases} \int_K 1_{(x_1, x_2]}(y) \mu_f(dy), & x_1 \leq x_2, \\ -\int_K 1_{(x_2, x_1]}(y) \mu_f(dy), & x_1 > x_2, \end{cases}$$

so that

$$|f(x_t) - f(x_s)| = \int_K 1_{(x_s, x_t]}(y) \mu_f(dy) + \int_K 1_{(x_t, x_s]}(y) \mu_f(dy).$$

Observe that $[1_{(x_s, x_t]}(y) + 1_{(x_t, x_s]}(y)]^p = 1_{(x_s, x_t]}(y) + 1_{(x_t, x_s]}(y)$, so that by Jensen's inequality,

$$|f(x_t) - f(x_s)|^p \leq \mu(K)^{p-1} \left\{ \int_K 1_{(x_s, x_t]}(y) \mu_f(dy) + \int_K 1_{(x_t, x_s]}(y) \mu_f(dy) \right\}.$$

Hence the Gagliardo seminorm of $f \circ X$ can be bounded by

$$[f \circ X]_{\theta, p}^p \leq 2\mu(K)^{p-1} \int_0^T \int_0^t \int_K \frac{1_{(x_s, x_t]}(y) + 1_{(x_t, x_s]}(y)}{(t - s)^{1+\theta p}} ds dt \mu_f(dy)$$

Because $\text{Leb} \otimes \mu_f(G_x) = 0$, it follows that the set

$$\tilde{A} = \{(s, t, y) \in [0, T]^2 \times K : x_s = y \text{ or } x_t = y\}$$

has $\text{Leb}^2 \otimes \mu_f$ -measure zero. Because

$$\frac{1_{(x_s, x_t]}(y) + 1_{(x_t, x_s]}(y)}{(t-s)^{1+\theta p}} = \frac{1_{(x_s, x_t)}(y) + 1_{(x_t, x_s)}(y)}{(t-s)^{1+\theta p}}$$

for all (s, t, y) outside the set \tilde{A} , we conclude that

$$[f \circ X]_{\theta, p}^p \leq 2\mu(K)^{p-1} \left\{ \int_K J_1(y) \mu_f(dy) + \int_K J_2(y) \mu_f(dy) \right\}, \quad (4.7)$$

where

$$J_1(y) = \int_0^T \int_0^t \frac{1_{(x_s, x_t)}(y)}{(t-s)^{1+\theta p}} ds dt \quad \text{and} \quad J_2(y) = \int_0^T \int_0^t \frac{1_{(x_t, x_s)}(y)}{(t-s)^{1+\theta p}} ds dt.$$

To estimate the above integrals, let us first rewrite $J_1(y)$ as an integral $J_1(y) = \int_0^T J_{1,t}(y) 1_{(x_t > y)} dt$, where

$$J_{1,t}(y) = \int_0^t \frac{1_{\{x_s < y\}}}{(t-s)^{1+\theta p}} ds.$$

By Lemma 4.5,

$$J_{1,t}(y) \leq (\theta p)^{-1} [x]_{\alpha, \infty}^{\theta p/\alpha} |x_t - y|^{-\theta p/\alpha}$$

for all t such that $x_t > y$. Hence

$$J_1(y) \leq (\theta p)^{-1} [x]_{\alpha, \infty}^{\theta p/\alpha} \int_0^T |x_t - y|^{-\theta p/\alpha} dt.$$

By applying the same argument to the function $t \mapsto -x_t$ at the point $-y$, we may also conclude that

$$J_{2,t}(y) := \int_0^t \frac{1_{\{x_s > y\}}}{(t-s)^{1+\theta p}} ds \leq (\theta p)^{-1} [x]_{\alpha, \infty}^{\theta p/\alpha} |x_t - y|^{-\theta p/\alpha}$$

for all t such that $x_t < y$. Therefore

$$J_2(y) = \int_0^T J_{2,t}(y) 1_{(x_t < y)} dt \leq (\theta p)^{-1} [x]_{\alpha, \infty}^{\theta p/\alpha} \int_0^T |x_t - y|^{-\theta p/\alpha} dt.$$

By combining these inequalities with (4.7), we obtain (4.6) for an increasing right-continuous f .

(ii) Let us now consider the general case with a right-continuous function f of finite variation. Then we can represent f as

$$f(t) = f(0) + f^+(t) - f^-(t),$$

where f^+ and f^- are the positive and negative variation functions of f defined by

$$f^\pm(t) = \begin{cases} V^\pm(0, t), & t \geq 0, \\ -V^\pm(t, 0), & t < 0. \end{cases}$$

The functions f^+ and f^- are increasing, bounded, and right-continuous [Kal02, Prop. 2.19], and the variation measure of f can be represented as $\mu_f = \mu_{f^+} + \mu_{f^-}$. Because

$$[f \circ x]_{\theta,p}^p \leq 2^{p-1} \left([f^+ \circ x]_{\theta,p}^p + [f^- \circ x]_{\theta,p}^p \right),$$

the claim follows by part (i). \square

Proof of Proposition 4.3. Because f is of locally finite variation, its set of discontinuities J_f is countable. Because X_t admits a probability density function for almost every t , it follows that

$$E \text{Leb}\{t : X_t \in J_f\} = E \int_0^T 1_{J_f}(X_t) dt = \int_0^T P(X_t \in J_f) dt = 0,$$

so that the set $\{t \in [0, T] : X_t \in J_f\}$ has Lebesgue measure zero with probability one. Hence by Lemma 4.4, $[f \circ X]_{\theta,p} = [f_r \circ X]_{\theta,p}$ almost surely, where f_r denotes the right-continuous version of f . Hence in what follows, we may and will assume that f is right-continuous.

Proposition 4.6 implies that $[f \circ X]_{\theta,p}^p \leq 2^p(\theta p)^{-1} [X]_{\alpha,\infty}^{\theta p/\alpha} Q$, where

$$Q = \mu_f(K_X)^{p-1} \int_0^T \int_{K_X} |X_t - y|^{-\theta p/\alpha} \mu_f(dy) dt,$$

and K_X is the closure of the range of X . Because the path of X is α -Hölder continuous almost surely, it suffices to verify that $Q < \infty$ almost surely. This is what we shall do next.

On the event $A_k = \{\|X\|_\infty \leq k\}$, the random variable Q is bounded from above by

$$Q_k = \mu_f([-k, k])^{p-1} \int_0^T \int_{[-k, k]} |X_t - y|^{-\theta p/\alpha} \mu_f(dy) dt.$$

Now

$$EQ_k \leq V_f(-k, k)^p \sup_{y \in \mathbb{R}} E \int_0^T |X_t - y|^{-\theta p/\alpha} dt.$$

where $V_f(-k, k) = \mu_f([-k, k])$ is the variation of f on $[-k, k]$. We know (Lemma D.1) that

$$E|X_t - y|^{-\theta p/\alpha} \leq 1 + \frac{2}{1 - \theta p/\alpha} \sup_x p_t(x) \leq 1 + \frac{2}{1 - \theta p/\alpha} \hat{p}_t$$

for all $y \in \mathbb{R}$ and for almost every t . By integrating the left and the right side of the above inequality over t , we find that

$$EQ_k \leq V_f(-k, k)^p \left(T + \frac{2}{1 - \theta p/\alpha} \int_0^T \hat{p}_t dt \right).$$

Because f has locally finite variation, it follows that $EQ_k < \infty$, and hence Q_k is finite almost surely.

We may now conclude that for all $k \geq 1$,

$$P(Q = \infty, A_k) \leq P(Q_k = \infty, A_k) \leq P(Q_k = \infty) = 0.$$

The almost sure continuity of X implies that $\|X\|_\infty$ is finite almost surely, and therefore $P(\cup_k A_k) = 1$. Hence

$$P(Q = \infty) = P(Q = \infty, \cup_k A_k) \leq \sum_k P(Q = \infty, A_k) = 0.$$

□

4.4 Proof of Theorem 3.1

Proof of Theorem 3.1. Choose some $\theta \in (1 - \beta, \alpha)$. By Proposition 4.1, it suffices to show that $f \circ X \in W_1^\theta(0+)$ and $Y \in W_\infty^{1-\theta}(T-)$ with probability one. Observe first that $f \circ X(0+)$ exists almost surely assumption. Moreover, by Lemma 4.2,

$$\|f \circ X\|_{W_1^\theta(0+)} \leq c(\|f \circ X\|_\infty + [f \circ X]_{\theta,1})$$

for $c = (1 - \theta)^{-1}T^{1-\theta} + \frac{1}{2}$. The term $\|f \circ X\|_\infty$ is finite almost surely, because X is continuous and f is locally bounded. The Gagliardo seminorm $[f \circ X]_{\theta,1}$ is almost surely finite by Proposition 4.3. Hence $f \circ X \in W_1^\theta(0+)$ almost surely.

Because Y is almost surely Hölder continuous of order $\beta > 1 - \theta$, we see that $\|Y\|_{W_\infty^{1-\theta}(T-)}$ is finite by Lemma 4.2. The Hölder continuity of Y also implies that $Y(T-)$ exists almost surely. Therefore $Y \in W_\infty^{1-\theta}(T-)$ with probability one, and the claim now follows by Proposition 4.1.

□

4.5 Proof of Theorem 3.4

Assume without loss of generality that $t = T$. Denote by $\phi = f'$ the derivative of f . We define a smooth approximation ϕ_n of ϕ as in (A.1). Define also a smooth approximation f_n of f by the same formula. Then f_n is infinitely differentiable with derivative $f'_n = \phi_n$. Because ϕ_n is Lipschitz continuous, we see that $\phi_n \circ X$ is Hölder continuous of order $\alpha > 1/2$, and it follows [Zäh98, Theorem 4.3.1] that

$$f_n(X_T) - f_n(X_0) = (\text{ZS}) \int_0^T \phi_n(X_t) dX_t \quad (4.8)$$

almost surely.

We will next verify that (4.8) remains valid in the limit as $n \rightarrow \infty$. By Proposition A.1, $[\phi_n \circ X - \phi \circ X]_{\theta,1} \xrightarrow{P} 0$. Observe next that the definition of ϕ_n guarantees that for any $c > 0$,

$$\sup_{|x| \leq c} |\phi_n(x) - \phi(x)| \leq \sup_{|x| \leq c+1} |\phi(x)| + \sup_{|x| \leq c} |\phi(x)|$$

for all n . Therefore,

$$t^{-\theta} |\phi_n(X_t) - \phi(X_t)| \leq 2t^{-\theta} \sup_{|x| \leq \|X\|_\infty + 1} |\phi(x)|,$$

where $\|X\|_\infty = \sup_{t \in [0, T]} |X_t|$ is almost surely finite. Because $\phi_n \rightarrow \phi$ pointwise and the right side above is almost surely integrable with respect to the Lebesgue measure on $[0, T]$, we conclude by dominated convergence that

$$\int_0^T \frac{|\phi_n(X_t) - \phi(X_t)|}{t^\theta} dt \rightarrow 0$$

almost surely. Our conclusion is that $\|\phi_n \circ X - \phi \circ X\|_{W_1^\theta(0+)} \rightarrow 0$ almost surely. Hence by Proposition 4.1, the right side of (4.8) converges according to

$$(\text{ZS}) \int_0^T \phi_n(X_t) dX_t \xrightarrow{P} (\text{ZS}) \int_0^T \phi(X_t) dX_t.$$

On the other hand, the left side of (4.8) converges almost surely to $f(X_T) - f(X_0)$. Because the limits must agree with probability one, we conclude that

$$f(X_T) - f(X_0) = (\text{ZS}) \int_0^T \phi(X_t) dX_t$$

almost surely, and the proof of Theorem 3.4 is complete.

4.6 Riemann–Stieltjes integrability

In this section we will prove Theorem 3.5 and Theorem 3.9. Let X and Y be Hölder continuous random processes on $[0, T]$ and f a function of locally finite variation fulfilling the assumptions of Theorem 3.5. Fix a tagged partition $\pi = (t_0, t_1, \dots, t_k; \xi_1, \dots, \xi_k)$ such that $0 = t_0 < t_1 < \dots < t_k$ and $\xi_i \in [t_{i-1}, t_i]$. Because Y is continuous on $[0, T]$, we see that (cf. [Zäh98, Proposition 2.2])

$$Y_{t_i} - Y_{t_{i-1}} = (\text{ZS}) \int_0^T 1_{(t_{i-1}, t_i]}(t) dY_t$$

for all i . Hence the Riemann–Stieltjes sum of $f \circ X$ against Y along the tagged partition π can be written as

$$\text{RS}(f \circ X, Y, \pi) = (\text{ZS}) \int_0^T \sum_{i=1}^k f(X_{\xi_i}) 1_{(t_{i-1}, t_i]}(t) dY_t = (\text{ZS}) \int_0^T f(\hat{X}_\pi(t)) dY_t,$$

where a π -interpolated version of X is defined by

$$\hat{X}_\pi(t) = \sum_{i=1}^k X_{\xi_i} 1_{(t_{i-1}, t_i]}(t)$$

By Proposition 4.1, we find that the approximation error

$$\Delta(\pi) = \text{RS}(f \circ X, g, \pi) - (\text{ZS}) \int_0^T f(X(t)) dY(t) \quad (4.9)$$

is bounded by

$$|\Delta(\pi)| = \left| (\text{ZS}) \int_0^T (f(\hat{X}_\pi(t)) - f(X_t)) dY_t \right| \leq \frac{\|h_\pi\|_{W_1^\theta(0+)} \|Y\|_{W_\infty^{1-\theta}(T-)}}{\Gamma(\theta)\Gamma(1-\theta)}, \quad (4.10)$$

where

$$h_\pi(t) = f \circ \hat{X}_\pi(t) - f \circ X(t).$$

denotes the interpolation error at t .

Proof of Theorem 3.5. Suppose that the assumptions of Theorem 3.5 prevail with $\alpha, \beta \in (0, 1)$ and $\delta, \lambda > 1$ such that $\alpha + \beta > 1$ and $\frac{\lambda}{1+\delta} + \beta > 1$. Consider an arbitrary $0 < a < \frac{\lambda}{1+\delta} + \beta - 1$.

Let us choose $\theta \in (0, 1)$ such that $1 - \beta < \theta < \alpha$ and $\theta < \frac{\lambda}{1+\delta} - a$. Observe that by Lemma 4.2,

$$\|Y\|_{W_\infty^{1-\theta}(T-)} \leq c_1[Y]_{\beta, \infty},$$

where $c_1 = (1 + \epsilon_1^{-1})T^{\epsilon_1}$ with $\epsilon_1 = \beta - (1 - \theta)$, so that by (4.10), the error term in (4.9) is bounded by

$$|\Delta(\pi)| \leq c_2 \|h_\pi\|_{W_1^\theta(0+)} [Y]_{\beta, \infty}$$

for $c_2 = c_1 \Gamma(\theta)^{-1} \Gamma(1 - \theta)^{-1}$.

(i) Assume first that f is right-continuous and has finite variation. Now by Lemma 4.7, $E\|h_\pi\|_{W_1^\theta(0+)} \leq cd_1(f)\|\pi\|^\rho$ for $\rho = \frac{\lambda}{1+\delta} - \theta > 0$. Hence $\|h_\pi\|_{W_1^\theta(0+)} \xrightarrow{P} 0$ as $\|\pi\| \rightarrow 0$. Because $[Y]_{\beta, \infty}$ is finite almost surely, we conclude that $|\Delta(\pi)| \xrightarrow{P} 0$ as $\|\pi\| \rightarrow 0$.

(ii) Assume next that f is right-continuous with locally, but not globally, finite variation. Let $A_k = \{\|X\|_\infty \leq k\}$. On this event we may replace f by its truncated version $f^{(k)}$ which has finite variation. On this event $h_\pi = h_\pi^{(k)}$ where the latter random function corresponds to replacing f by $f^{(k)}$. By the first part, $\|h_\pi^{(k)}\|_{W_1^\theta(0+)} \xrightarrow{P} 0$. Because $h_\pi^{(k)} = h_\pi$ on the event A_k , we conclude that $1_{A_k} \|h_\pi^{(k)}\|_{W_1^\theta(0+)} \xrightarrow{P} 0$ for all $k \geq 1$. Because $P(\cup_k A_k) = P(\|X\|_\infty < \infty) = 1$, we conclude (Lemma C.1) that $\|h_\pi\|_{W_1^\theta(0+)} \xrightarrow{P} 0$ also in this case.

(iii) Assume now that

$$\sum_{n=1}^{\infty} \|\pi_n\|^a < \infty$$

for some $a < \frac{\lambda}{1+\delta} + \beta - 1$. We may and will assume that this a is the same a that was fixed in the beginning of the proof. Then for any right-continuous

function of finite variation, by applying Lemma 4.7 we see that

$$E \left(\sum_{n \geq 1} \|h_n\|_{W_1^\theta(0+)} \right) \leq \sum_{n \geq 1} c \|\pi\|^{\frac{\lambda}{1+\delta}-\theta} \leq c \sum_{n \geq 1} \|\pi\|^a < \infty.$$

Hence $\|h_n\|_{W_1^\theta(0+)} \rightarrow 0$ almost surely, and hence also $|\Delta_n| \rightarrow 0$ almost surely. Making the same conclusion for any right-continuous f of locally finite variation can be done as in (ii). \square

Proof of Theorem 3.9. Suppose that the assumptions of Theorem 3.5 prevail with $\alpha, \beta \in (0, 1)$ and $\delta, \lambda > 1$ such that $\alpha + \beta > 1$ and $\frac{\lambda}{1+\delta} + \beta > 1$. Consider an arbitrary $0 < a < \frac{\lambda}{1+\delta} + \beta - 1$.

Let us first choose $\theta \in (0, 1)$ such that $1 - \beta < \theta < \alpha$ and $\theta < \frac{\lambda}{1+\delta} - a$, and then a small enough $p \in (1, 2)$ so that

$$p \leq \min \left(\delta, 1 + \frac{\lambda}{1+\delta}, \frac{\frac{\lambda}{1+\delta} - \theta + 1}{a + 1} \right).$$

Observe that by Lemma 4.2,

$$\|Y\|_{W_\infty^{1-\theta}(T-)} \leq c_1[Y]_{\beta, \infty},$$

where $c_1 = (1 + \epsilon_1^{-1})T^{\epsilon_1}$ with $\epsilon_1 = \beta - (1 - \theta)$, so that by (4.10), the Riemann–Stieltjes integration error is bounded by

$$|\Delta(\pi)| \leq c_2 \|h_\pi\|_{W_1^\theta(0+)} [Y]_{\beta, \infty}$$

for $c_2 = c_1 \Gamma(\theta)^{-1} \Gamma(1 - \theta)^{-1}$. If f has finite total variation, then we find by Lemma 4.7 and our choice of p that (for $\|\pi\| \leq 1$)

$$\left(E \|h_\pi\|_{W_1^\theta(0+)}^p \right)^{1/p} \leq c_3 d_1(f)^{1/p} \|\pi\|^{(1 + \frac{\lambda}{1+\delta} - \theta - p)/p} \leq c_3 d_1(f)^{1/p} \|\pi\|^a, \quad (4.11)$$

so that for $1/q = 1 - 1/p$,

$$E |\Delta(\pi)| \leq c_2 \left(E \|h_\pi\|_{W_1^\theta(0+)}^p \right)^{1/p} \left(E [Y]_{\beta, \infty}^q \right)^{1/q} \leq c_2 c_3 d_1(f)^{1/p} \left(E [Y]_{\beta, \infty}^q \right)^{1/q} \|\pi\|^a,$$

where

$$d_1(f) = \begin{cases} V(f)^p, & \epsilon = 0, \\ V(f)^p + (\sup_{|x| < \epsilon} |f'(x)|)^p, & \epsilon > 0. \end{cases}$$

This proves the claim under the assumption that f has finite total variation.

Assume next that the total variation of f is infinite. Denote by A_k the event that $k - 1 < \|X\|_\infty \leq k$. On this event, we see that

$$h_\pi = f \circ \hat{X}_\pi - f \circ X = f^{(k)} \circ \hat{X}_\pi - f^{(k)} \circ X =: h_\pi^{(k)},$$

where $f^{(k)}$ is the truncated version of f defined by

$$f^{(k)}(x) = \begin{cases} f(-k), & x < -k, \\ f(x), & |x| \leq k, \\ f(k), & x > k. \end{cases}$$

By the same argument as above,

$$|\Delta(\pi)|1_{A_k} \leq c_2 \|h_\pi^{(k)}\|_{W_1^\theta(0+)} [Y]_{\beta, \infty} 1_{A_k}.$$

Then we may choose a large enough $q \in (1, \infty)$ so that $1/p + 2/q = 1$. Then by Hölder's inequality for three random variables,

$$E(|\Delta(\pi)|1_{A_k}) \leq c_2 \left(E \|h_\pi^{(k)}\|_{W_1^\theta(0+)}^p \right)^{1/p} \left(E \|Y\|_{\beta, \infty}^q \right)^{1/q} P(A_k)^{1/q}.$$

By applying Lemma 4.7 again, we see that (4.11) is valid with f replaced by $f^{(k)}$. Therefore,

$$E|\Delta(\pi)| = E \sum_{k=0}^{\infty} |\Delta(\pi)|1_{A_k} \leq c_2 c_3 \left(E \|Y\|_{\beta, \infty}^q \right)^{1/q} \left(\sum_{k=1}^{\infty} d_1(f^{(k)})^{1/p} P(A_k)^{1/q} \right) \|\pi\|^a.$$

Now because $f^{(k)}(x) = f(x)$ for $|x| \leq k$,

$$d_1(f^{(k)}) = \begin{cases} V(f^{(k)})^p, & \epsilon = 0, \\ V(f^{(k)})^p + (\sup_{|x| < \epsilon} |f'(x)|)^p, & \epsilon > 0. \end{cases}$$

Because f has infinite total variation, $V(f^{(k)}) = V_{[-k, k]}(f) \rightarrow \infty$, and it follows that

$$d_1(f^{(k)})^{1/p} \leq 2^{1/p} V(f^{(k)}) = 2^{1/p} V_{[-k, k]}(f)$$

for all large values of k . Because $P(A_k) \leq P(\|X\|_\infty > k - 1)$, we find that

$$\left(\sum_{k=1}^{\infty} d_1(f^{(k)})^{1/p} P(A_k)^{1/q} \right) < \infty,$$

which confirms the validity of the claim. \square

Lemma 4.7. *Let f be a right-continuous function of finite variation. Let $X : [0, T] \rightarrow \mathbb{R}$ be a random process such that X_t admits a probability density function p_t bounded according to Assumption 2.1, and that*

$$c_2 = \sup_{0 \leq s < t \leq T} \frac{E|X_t - X_s|^\delta}{|t - s|^\lambda} < \infty$$

for some $\delta \geq 1$ and $\lambda > 1$. Assume also that there exist $\epsilon \geq 0$ and $c_0 < \infty$ such that

- $p_t(x) \leq c_0$ for all $|x| \geq \epsilon/2$ and all $t \in (0, T)$, and

- f restricted to $(-\epsilon, \epsilon)$ is Lipschitz continuous.

Then for any $1 \leq p \leq \delta$ and $0 < \theta < \min(2 - p, 1 + \frac{\lambda}{1+\delta} - p)$, there exists a finite constant c such that the interpolation error $h_\pi = f \circ \hat{X}_\pi - f \circ X$ satisfies

$$E \|h_\pi\|_{W_1^\theta(0+)}^p \leq c d_1(f) \|\pi\|^{1 + \frac{\lambda}{1+\delta} - p - \theta}$$

for any tagged partition π , where

$$d_1(f) = \begin{cases} V(f)^p, & \epsilon = 0, \\ V(f)^p + (\sup_{|x| < \epsilon} |f'(x)|)^p, & \epsilon > 0, \end{cases}$$

Proof. By Proposition B.1, we see that there exists a constant c_6 such that

$$E |f(X_t) - f(X_s)|^p \leq c_6 d_1(f) |t - s|^r, \quad (4.12)$$

for all $s, t \in [0, T]$, where $r = \frac{\lambda}{1+\delta}$.

Now consider a tagged partition π , and denote the intervals of the partition by $A_i = (t_{i-1}, t_i]$. By applying Jensen's inequality we see that

$$\begin{aligned} \|h_\pi\|_{W_1^\theta(0+)}^p &= \left(\int_0^T |h_\pi(t)| t^{-\theta} dt + \frac{1}{2} \int_0^T \int_0^T \frac{|h_\pi(t) - h_\pi(s)|}{|t - s|^{1+\theta}} ds dt \right)^p \\ &\leq c_3 \int_0^T |h_\pi(t)|^p t^{-\theta} dt + c_4 \int_0^T \int_0^T \frac{|h_\pi(t) - h_\pi(s)|^p}{|t - s|^{p+\theta}} ds dt. \end{aligned} \quad (4.13)$$

where $c_3 = 2^{p-1} (\int_0^T t^{-\theta} dt)^{p-1}$ and $c_4 = 2^{-1} (\int_0^T \int_0^T |t - s|^{-\theta} ds dt)^{p-1}$.

To estimate the first term on the right side of (4.13), observe first that

$$|h_\pi(t)|^p = \sum_i |f(X_{\xi_i}) - f(X_t)|^p 1_{(t_{i-1}, t_i]}(t).$$

Because $|t - \xi_i| \leq t_i - t_{i-1}$ for $t, \xi_i \in [t_{i-1}, t_i]$, we see by applying (4.12), the subadditivity of $t \mapsto t^{1-\theta}$, and the formula $\sum_{i=1}^k (t_i - t_{i-1}) = T$ that

$$\begin{aligned} E \int_0^T |h_\pi(t)|^p t^{-\theta} dt &= \sum_i \int_{t_{i-1}}^{t_i} E |f(X_{\xi_i}) - f(X_t)|^p t^{-\theta} dt \\ &\leq c_6 d_1(f) \sum_i (t_i - t_{i-1})^r \int_{t_{i-1}}^{t_i} t^{-\theta} dt \\ &= c_6 d_1(f) (1 - \theta)^{-1} \sum_i (t_i - t_{i-1})^r (t_i^{1-\theta} - t_{i-1}^{1-\theta}) \\ &\leq c_6 d_1(f) (1 - \theta)^{-1} \sum_i (t_i - t_{i-1})^{r-\theta+1} \\ &\leq c_6 d_1(f) T (1 - \theta)^{-1} \|\pi\|^{r-\theta}. \end{aligned}$$

We conclude that the expectation of the first term on the right side of (4.13) is bounded by

$$E \int_0^T |h_\pi(t)|^p t^{-\theta} dt \leq c_6 d_1(f) T (1 - \theta)^{-1} \|\pi\|^{r-\theta}. \quad (4.14)$$

Let us next analyze the second term in the upper bound of (4.13). By computing integrals part by part along the partition π , we see that

$$E \left(\int_0^T \int_0^T \frac{|h_\pi(t) - h_\pi(s)|^p}{|t - s|^{p+\theta}} \right) ds dt = \sum_i m_{i,i} + 2 \sum_{i,j:i < j} m_{i,j},$$

where

$$m_{i,j} = E \int_{A_i} \left(\int_{A_j} \frac{|f(X_{\xi_i}) - f(X_t) - f(X_{\xi_j}) + f(X_s)|^p}{|t - s|^{p+\theta}} dt \right) ds.$$

(i) For the diagonal terms, we see by (4.12) that

$$\begin{aligned} m_{i,i} &= \int_{A_i} \int_{A_i} \frac{E|f(X_t) - f(X_s)|^p}{|t - s|^{p+\theta}} ds dt = \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^t \frac{E|f(X_t) - f(X_s)|^p}{(t - s)^{p+\theta}} ds dt \\ &\leq c_6 d_1(f) \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^t (t - s)^{r-\theta-p} ds dt. \end{aligned}$$

Denote $\rho = r - \theta - p + 1$. Then $\rho > 0$ and

$$\int_{t_{i-1}}^t (t - s)^{\rho-1} ds dt = \rho^{-1}(1+\rho)^{-1}(t_i - t_{i-1})^{1+\rho} \leq \rho^{-1}(1+\rho)^{-1} \|\pi\|^\rho (t_i - t_{i-1}),$$

so that

$$m_{i,i} \leq c_6 d_1(f) \rho^{-1} (1 + \rho)^{-1} \|\pi\|^\rho (t_i - t_{i-1}).$$

By summing over i we find that

$$\sum_i m_{i,i} \leq c'_5 d_1(f) T \|\pi\|^\rho$$

with $c'_5 = \rho^{-1}(1 + \rho)^{-1} c_6$.

(ii) For the offdiagonal terms with $i < j$, we find that $m_{i,j} \leq m_{i,j}(1) + m_{i,j}(2)$ where

$$m_{i,j}(1) = \int_{A_i} \left(\int_{A_j} \frac{E|f(X_{\xi_i}) - f(X_s)|^p}{|t - s|^{p+\theta}} dt \right) ds, \quad m_{i,j}(2) = \int_{A_i} \left(\int_{A_j} \frac{E|f(X_{\xi_j}) - f(X_t)|^p}{|t - s|^{p+\theta}} dt \right) ds.$$

For the first term, for $i < j$, we find by (4.12) that

$$m_{i,j}(1) \leq c_6 d_1(f) \|\pi\|^r \int_{A_i} \left(\int_{A_j} |t - s|^{-\theta-p} dt \right) ds.$$

By summing over $j > i$, and then over i , we find that

$$\sum_i \sum_{j>i} m_{i,j}(1) \leq c_6 d_1(f) \|\pi\|^r \sum_i \int_{A_i} \left(\sum_{j>i} \int_{A_j} (t - s)^{-\theta-p} dt \right) ds.$$

Observe next that

$$\begin{aligned}
\int_{A_i} \left(\sum_{j>i} \int_{A_j} (t-s)^{-\theta-p} dt \right) ds &= \int_{t_{i-1}}^{t_i} \left(\int_{t_i}^T (t-s)^{-\theta-p} dt \right) ds \\
&\leq \int_{t_{i-1}}^{t_i} \left(\int_{t_i}^{\infty} (t-s)^{-\theta-p} dt \right) ds \\
&= \int_{t_{i-1}}^{t_i} (\theta+p-1)^{-1} (t_i-s)^{1-\theta-p} ds \\
&= (\theta+p-1)^{-1} (2-(\theta+p))^{-1} (t_i-t_{i-1})^{2-(\theta+p)}.
\end{aligned}$$

Hence

$$\begin{aligned}
\sum_i \sum_{j>i} m_{i,j}(1) &\leq c_5'' d_1(f) \|\pi\|^r \sum_i (t_i - t_{i-1})^{2-(\theta+p)} \\
&\leq c_5'' d_1(f) T \|\pi\|^{1+r-\theta-p} \\
&= c_5'' d_1(f) T \|\pi\|^\rho,
\end{aligned}$$

where $c_5'' = (\theta+p-1)^{-1} (2-(\theta+p))^{-1} c_6$.

For the second term, for $i < j$, we find by (4.12) in the same way as above that

$$m_{i,j}(2) \leq c_6 \|\pi\|^r \int_{A_i} \left(\int_{A_j} |t-s|^{-\theta-p} dt \right) ds.$$

This upper bound is the same as that earlier, with i and j interchanged. Therefore $\sum_j \sum_{i:i<j} m_{i,j}(2)$ satisfies the same upper bound as above, and we conclude that

$$\begin{aligned}
E \left(\int_0^T \int_0^T \frac{|h_\pi(t) - h_\pi(s)|^p}{|t-s|^{p+\theta}} ds dt \right) &\leq \sum_i m_{i,i} + 4 \sum_i \sum_{j>i} m_{i,j}(1) \\
&\leq (c_5' + 4c_5'') T d_1(f) \|\pi\|^\rho.
\end{aligned} \tag{4.15}$$

By substituting (4.14) and (4.15) into (4.13), we conclude that

$$\|h_\pi\|_{W_1^\theta(0+)}^p \leq c_3 c_6 T (1-\theta)^{-1} d_1(f) \|\pi\|^{r-\theta} + c_4 (c_5' + 4c_5'') T d_1(f) \|\pi\|^\rho.$$

Because $r-\theta = \rho + (p-1) \geq \rho$, we conclude that $\|\pi\|^{r-\theta} \leq T \|\pi\|^\rho$ for $\|\pi\| \leq T$. Hence the claim follows. \square

A Approximation by smooth functions

The purpose of this section is to approximate a function f of locally finite variation by a smooth function f_n , and to show that for a sufficiently regular random process X , the approximation can be made so that $f_n \circ X$ is with high

probability close to $f \circ X$ in the Gagliardo seminorm. We approximate f in a standard manner by

$$f_n(x) = Ef(x - n^{-1}\xi) \quad (\text{A.1})$$

where ξ is a random number in $[0, 1]$ with an infinitely differentiable probability density function. The following is the main result of this section.

Proposition A.1. *Let X be a Hölder continuous random process of order $\alpha > 0$ with a probability density function bounded according to Assumption 2.1. Then for any right-continuous function f of locally finite variation and any $\theta \in (0, \alpha)$,*

$$[f_n \circ X - f \circ X]_{\theta,1} \xrightarrow{P} 0.$$

The proof of the proposition utilizes the following two auxiliary results.

Lemma A.2. *Let f be bounded, increasing, and right continuous. Then the function f_n defined by (A.1) is increasing, infinitely differentiable and bounded, and $f_n \rightarrow f$ pointwise. Moreover, the Lebesgue–Stieltjes measure μ_{f_n} of f_n satisfies*

$$\int_{\mathbb{R}} h(y) \mu_{f_n}(dy) \rightarrow \int_{\mathbb{R}} h(y) \mu_f(dy)$$

for all bounded continuous h on \mathbb{R} .

Proof. The total mass of μ_f equals $\mu(\mathbb{R}) = f(+\infty) - f(-\infty)$ and is finite and nonnegative. Lebesgue’s dominated convergence theorem implies that $f_n \rightarrow f$ pointwise. Lebesgue’s monotone convergence theorem implies that $f_n(\pm\infty) = f(\pm\infty)$.

We can write f_n as a convolution

$$f_n(x) = \int_{-\infty}^{\infty} f(x-y)\phi_n(y) dy$$

where $\phi_n(y) = n\phi(ny)$ is the probability density function of $n^{-1}\xi$. This implies that f_n is infinitely differentiable. The definition also shows that f_n is increasing. Hence the Lebesgue–Stieltjes measure of f_n equals

$$\mu_{f_n}(dx) = f'_n(x)dx.$$

Because $f_n(\pm\infty) = f(\pm\infty)$, we see that $\mu_n(\mathbb{R}) = \mu(\mathbb{R})$. The pointwise convergence $f_n \rightarrow f$ implies that

$$\mu_n(-\infty, x] = f_n(x) - f_n(-\infty) = f_n(x) - f(-\infty) \rightarrow f(x) - f(-\infty) = \mu(-\infty, x]$$

for all real numbers x . Assuming that f is not constant, we have $\mu(\mathbb{R}) \in (0, \infty)$, and we find by applying the portmanteau theorem that $\mu_n(dx)/\mu(\mathbb{R}) \rightarrow \mu(dx)/\mu(\mathbb{R})$ weakly. This implies the claim. If f is constant, the claim is trivial because then $\mu_f = \mu_{f_n} = 0$. \square

Lemma A.3. *If X is a Hölder continuous random process of order $\alpha > 0$ with a probability density function bounded according to Assumption 2.1, then for any $\theta \in (0, \alpha)$ and $q \geq \theta/\alpha$, the functions*

$$y \mapsto E \left\{ (1 + [X]_{\alpha, \infty})^{-q} \int_0^T \int_0^t \frac{1_{\{X_s < y < X_t\}}}{(t-s)^{1+\theta}} ds dt \right\}$$

and

$$y \mapsto E \left\{ (1 + [X]_{\alpha, \infty})^{-q} \int_0^T \int_0^t \frac{1_{\{X_t < y < X_s\}}}{(t-s)^{1+\theta}} ds dt \right\}$$

are bounded and continuous.

Proof. (i) Denote $\phi(y) = E\Phi(y)$, where

$$\Phi(y) = (1 + [X]_{\alpha, \infty})^{-q} \int_0^T \int_0^t \frac{1_{(X_s, X_t)}(y)}{(t-s)^{1+\theta}} ds dt.$$

By Lemma 4.5,

$$\Phi(y) \leq \theta^{-1} \frac{[X]_{\alpha, \infty}^{\theta/\alpha}}{(1 + [X]_{\alpha, \infty})^q} \int_0^T |X_t - y|^{-\theta/\alpha} dt \leq \theta^{-1} \int_0^T |X_t - y|^{-\theta/\alpha} dt.$$

Hence by Lemma D.1 and Assumption 2.1, we see that

$$\phi(y) = E\Phi(y) \leq \theta^{-1} \int_0^T \left(1 + \frac{2}{1 - \theta/\alpha} \hat{p}_t \right) dt.$$

The right side above is finite and does not depend on y . Hence the function ϕ is bounded.

(ii) Let us next verify that ϕ is right-continuous at $y \in \mathbb{R}$. Fix $\epsilon > 0$ and observe that

$$1_{\{X_s < y + \epsilon < X_t\}} - 1_{\{X_s < y < X_t\}} = 1_{\{y \leq X_s < y + \epsilon < X_t\}} - 1_{\{X_s < y < X_t \leq y + \epsilon\}}.$$

Therefore,

$$\phi(y + \epsilon) - \phi(y) = E \int_0^T \int_0^t \Phi_{1, \epsilon}(s, t) ds dt - E \int_0^T \int_0^t \Phi_{2, \epsilon}(s, t) ds dt, \quad (\text{A.2})$$

where

$$\begin{aligned} \Phi_{1, \epsilon}(s, t) &= (1 + [X]_{\alpha, \infty})^{-q} \frac{1_{\{y \leq X_s < y + \epsilon < X_t\}}}{(t-s)^{1+\theta}}, \\ \Phi_{2, \epsilon}(s, t) &= (1 + [X]_{\alpha, \infty})^{-q} \frac{1_{\{X_s < y < X_t \leq y + \epsilon\}}}{(t-s)^{1+\theta}}. \end{aligned}$$

Assumption 2.1 implies that $1_{\{X_s(\omega)=y\}} = 0$ and $1_{\{X_t(\omega)=y\}} = 0$ for $P \otimes \text{Leb}^2$ -almost every (ω, s, t) in $\Omega \times [0, T]^2$ where $[0, T]_{<}^2 = \{(s, t) \in [0, T] : s < t\}$. Hence, as $\epsilon \rightarrow 0$,

$$\Phi_{1, \epsilon}(s, t) \rightarrow 0 \quad \text{and} \quad \Phi_{2, \epsilon}(s, t) \rightarrow 0.$$

for $P \otimes \text{Leb}^2$ -almost every (ω, s, t) .

Let us next verify that

$$E \int_0^T \int_0^t \Phi_{1,\epsilon}(s, t) ds dt \rightarrow 0 \quad (\text{A.3})$$

as $\epsilon \rightarrow 0$. To do this, choose $p \in (1, \frac{1+\alpha}{1+\theta})$ so small that $1/p \geq 1 + \theta - \alpha q$. Then

$$\Phi_{1,\epsilon}(s, t)^p = (1 + [X]_{\alpha, \infty})^{-\tilde{q}} \frac{1_{\{y \leq X_s < y + \epsilon < X_t\}}}{(t - s)^{1+\tilde{\theta}}} \leq (1 + [X]_{\alpha, \infty})^{-\tilde{q}} \frac{1_{\{X_s < y + \epsilon < X_t\}}}{(t - s)^{1+\tilde{\theta}}},$$

where $\tilde{q} = pq$ and $\tilde{\theta} = (1 + \theta)p - 1$. Our choice of p implies that $\tilde{q} \geq \tilde{\theta}/\alpha$ and $\tilde{\theta} \in (0, \alpha)$. By repeating the argument used in part (i) with q and θ replaced by \tilde{q} and $\tilde{\theta}$, we find that

$$E \int_0^T \int_0^t \Phi_{1,\epsilon}(s, t)^p ds dt \leq \tilde{\theta}^{-1} \int_0^T \left(1 + \frac{2}{1 - \tilde{\theta}/\alpha} \hat{p}_t\right) dt,$$

where the upper bound on the right is finite and does not depend on ϵ . We may hence conclude that the collection of functions $\Phi_{1,\epsilon}$ indexed by $\epsilon > 0$ is uniformly integrable with respect to $P \otimes \text{Leb}^2$, and we obtain (A.3).

Let us next verify that

$$E \int_0^T \int_0^t \Phi_{2,\epsilon}(s, t) ds dt \rightarrow 0 \quad (\text{A.4})$$

as $\epsilon \rightarrow 0$. This is not hard, because

$$\Phi_{2,\epsilon}(s, t) \leq (1 + [X]_{\alpha, \infty})^{-q} \frac{1_{\{X_s < y < X_t\}}}{(t - s)^{1+\theta}}$$

is valid for all (ω, s, t) . The right side is integrable with respect to $P \otimes \text{Leb}^2$, by the argument used in part (i). Hence (A.4) follows by Lebesgue's dominated convergence theorem. By substituting the limits (A.3)–(A.4) into the representation (A.2), we may conclude that $y \mapsto \phi(y)$ is right continuous.

(iii) Let us next verify that ϕ is left-continuous at $y \in \mathbb{R}$. Fix $\epsilon > 0$ and observe that

$$1_{\{X_s < y - \epsilon < X_t\}} - 1_{\{X_s < y < X_t\}} = 1_{\{X_s < y - \epsilon < X_t \leq y\}} - 1_{\{y - \epsilon \leq X_s < y < X_t\}}.$$

Therefore,

$$\phi(y - \epsilon) - \phi(y) = E \int_0^T \int_0^t \Psi_{1,\epsilon}(s, t) ds dt - E \int_0^T \int_0^t \Psi_{2,\epsilon}(s, t) ds dt, \quad (\text{A.5})$$

where

$$\begin{aligned} \Psi_{1,\epsilon}(s, t) &= (1 + [X]_{\alpha, \infty})^{-q} \frac{1_{\{X_s < y - \epsilon < X_t \leq y\}}}{(t - s)^{1+\theta}}, \\ \Psi_{2,\epsilon}(s, t) &= (1 + [X]_{\alpha, \infty})^{-q} \frac{1_{\{y - \epsilon \leq X_s < y < X_t\}}}{(t - s)^{1+\theta}}. \end{aligned}$$

By the same arguments as in part (ii) of the proof, we see that as $\epsilon \rightarrow 0$,

$$\Psi_{1,\epsilon}(s, t) \rightarrow 0 \quad \text{and} \quad \Psi_{2,\epsilon}(s, t) \rightarrow 0$$

for $P \otimes \text{Leb}^2$ -almost every (ω, s, t) . We also see that the collection of functions $\Psi_{1,\epsilon}$ indexed by $\epsilon > 0$ is uniformly integrable with respect to $P \otimes \text{Leb}^2$, and that $\Psi_{2,\epsilon}(s, t)$ is bounded from above by a $P \otimes \text{Leb}^2$ -integrable function. Hence,

$$E \int_0^T \int_0^t \Psi_{1,\epsilon}(s, t) ds dt \rightarrow 0 \quad \text{and} \quad E \int_0^T \int_0^t \Psi_{2,\epsilon}(s, t) ds dt \rightarrow 0,$$

and the left-continuity of ϕ follows from the representation (A.5).

(iv) Let us finally show that the function

$$\tilde{\phi}(y) = E \left\{ (1 + [X]_{\alpha, \infty})^{-q} \int_0^T \int_0^t \frac{1_{\{X_t < y < X_s\}}}{(t-s)^{1+\theta}} ds dt \right\}$$

is bounded and continuous. To do this, define another random process \tilde{X} by $\tilde{X}_t = -X_t$. Then \tilde{X} has α -Hölder continuous paths with $[\tilde{X}]_{\alpha, \infty} = [X]_{\alpha, \infty}$, and we may represent $\tilde{\phi}$ as

$$\tilde{\phi}(y) = E \left\{ (1 + [\tilde{X}]_{\alpha, \infty})^{-q} \int_0^T \int_0^t \frac{1_{\{\tilde{X}_s < -y < \tilde{X}_t\}}}{(t-s)^{1+\theta}} ds dt \right\}.$$

The boundedness and continuity of $\tilde{\phi}$ now follow from the boundedness and continuity of ϕ . \square

Proof of Proposition A.1. (i) Let us first assume that f right-continuous, increasing, and bounded. Let f_n be the approximation of f defined by (A.1). Denote

$$Z_n(s, t) = \frac{f_n(X_t) - f_n(X_s)}{(t-s)^{1+\theta}} \quad \text{and} \quad Z(s, t) = \frac{f(X_t) - f(X_s)}{(t-s)^{1+\theta}}.$$

Observe that

$$E \int_0^T \int_0^t \frac{|Z_n(s, t)|}{(1 + [X]_{\alpha, \infty})^q} ds dt = \int_{\mathbb{R}} h(y) \mu_{f_n}(dy)$$

where $\mu_{f_n}(dy) = f'_n(y)dy$ is the Lebesgue–Stieltjes measure of f_n , and

$$h(y) = E \left\{ (1 + [X]_{\alpha, \infty})^{-q} \int_0^T \int_0^t \frac{1_{\{X_s < y < X_t\}} + 1_{\{X_t < y < X_s\}}}{(t-s)^{1+\theta}} ds dt \right\}$$

is continuous by Lemma A.3. Hence, by Lemma A.2,

$$\int_{\mathbb{R}} h(y) \mu_{f_n}(dy) \rightarrow \int_{\mathbb{R}} h(y) \mu_f(dy),$$

from which we conclude that

$$E \int_0^T \int_0^t \frac{|Z_n(s, t)|}{(1 + [X]_{\alpha, \infty})^q} ds dt \rightarrow E \int_0^T \int_0^t \frac{|Z(s, t)|}{(1 + [X]_{\alpha, \infty})^q} ds dt.$$

This observation combined with the fact that $Z_n(s, t) \rightarrow Z(s, t)$ for $P \otimes \text{Leb}^2$ -almost every (ω, s, t) , allows us to conclude [Kal02, Lemma 1.32] that

$$E \int_0^T \int_0^t \frac{|Z_n(s, t) - Z(s, t)|}{(1 + [X]_{\alpha, \infty})^q} ds dt \rightarrow 0.$$

As a consequence,

$$(1 + [X]_{\alpha, \infty})^{-q} \int_0^T \int_0^t |Z_n(s, t) - Z(s, t)| ds dt \xrightarrow{P} 0,$$

and therefore also

$$[f_n \circ X - f \circ X]_{\theta, 1} = 2 \int_0^T \int_0^t |Z_n(s, t) - Z(s, t)| ds dt \xrightarrow{P} 0.$$

(ii) Let us next assume that f is right-continuous and of finite variation. Then we may write $f = f^{(1)} - f^{(2)}$ where $f^{(i)}$ are right-continuous, increasing, and bounded. The mollification of f can be written as $f_n = f_n^{(1)} - f_n^{(2)}$. Hence

$$\begin{aligned} [f_n \circ X - f \circ X]_{\theta, 1} &= [(f_n^{(1)} \circ X - f^{(1)} \circ X) + (f_n^{(2)} \circ X - f^{(2)} \circ X)]_{\theta, 1} \\ &\leq [(f_n^{(1)} \circ X - f^{(1)} \circ X)]_{\theta, 1} + [(f_n^{(2)} \circ X - f^{(2)} \circ X)]_{\theta, 1} \end{aligned}$$

By part (i), both terms on the right tend to zero in probability. Hence the claim is true also in this case.

(iii) Let us next assume that f is right-continuous and of locally finite variation. Let $A_m = \{\|X\|_\infty \leq m\}$ be the event that the path of X is bounded by m . Define $f^{(m)}$ by

$$f^{(m)}(y) = \begin{cases} f(-m), & y < -m, \\ f(y), & |y| \leq m, \\ f(m), & y > m. \end{cases}$$

Then $f^{(m)}$ is right-continuous and of finite variation. Moreover, on the event A_m , $f \circ X = f^{(m)} \circ X = f^{(m+1)} \circ X$ and $f_n \circ X = f_n^{(m+1)} \circ X$ for all $n \geq 1$. Hence, as a consequence of part (ii), we see that for any m and any $\epsilon > 0$,

$$\begin{aligned} P([f_n \circ X - f \circ X]_{\theta, 1} > \epsilon, A_m) &= P([f_n^{(m+1)} \circ X - f^{(m+1)} \circ X]_{\theta, 1} > \epsilon, A_m) \\ &\leq P([f_n^{(m+1)} \circ X - f^{(m+1)} \circ X]_{\theta, 1} > \epsilon) \\ &\rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$. We conclude that

$$1_{A_m} [f_n \circ X - f \circ X]_{\theta, 1} \xrightarrow{P} 0$$

for all $m \geq 1$. Because $P(\cup_m A_m) = P(\|X\|_\infty < \infty) = 1$, it follows (Lemma C.1) that $[f_n \circ X - f \circ X]_{\theta, 1} \xrightarrow{P} 0$, so the claim holds also in this case. \square

B Stochastic continuity of finite variation functions

The following result establishes a stochastic continuity property of a finite variation function f evaluated at two random inputs X_1 and X_2 which are close to each other in expectation, and such that the probability distribution of X_1 admits a density function p_1 . The result can be seen as an extension of [Avi09, Theorem 2.4] into a setting where p_1 is not globally bounded.

Below the requirement that f is right-continuous can be relaxed by similar arguments as in the proof of [Avi09, Theorem 2.4]. The result stated here suffices for the purposes of this paper.

Proposition B.1. *Let $1 \leq p \leq q < \infty$ and $\epsilon \geq 0$, and assume that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a right-continuous function of finite variation such that the restriction of f into $(-\epsilon, \epsilon)$ is Lipschitz continuous. Let X_1 be random variable which admits a probability density function $p_1(x)$. Then*

$$E|f(X_1) - f(X_2)|^p \leq c d_1(f)(1 + d_2(p_1)) \left(E|X_1 - X_2|^q \right)^{\frac{1}{1+q}}$$

for all random variables X_2 such that $E|X_1 - X_2|^q \leq 1$, where $c = 2^{p-1+q/(1+q)}(1+q)^{1/(1+q)}$,

$$d_1(f) = \begin{cases} V(f)^p, & \epsilon = 0, \\ V(f)^p + \left(\sup_{|x| < \epsilon} |f'(x)| \right)^p, & \epsilon > 0, \end{cases}$$

and

$$d_2(p_1) = \begin{cases} \left(\sup_{x \in \mathbb{R}} p_1(x) \right)^{\frac{q}{q+1}}, & \epsilon = 0, \\ \left(2\epsilon^{-1} + \sup_{|x| > \epsilon/2} p_1(x) \right)^{\frac{q}{q+1}}, & \epsilon > 0. \end{cases}$$

The proof of Proposition B.1 is based on the following two auxiliary results.

Lemma B.2. *Let X be a random number and let A, B be events defined on the same probability space as X .*

(i) *If ϕ is positive and increasing, then*

$$E\phi(X)1_{A \cap B} \geq E\phi(X)1_{A \cap \{X \leq b\}}$$

whenever $P(A, B) \geq P(A, X \leq b)$.

(ii) *If ϕ is positive and decreasing, then*

$$E\phi(X)1_{A \cap B} \geq E\phi(X)1_{A \cap \{X > b\}}$$

whenever $P(A, B) \geq P(A, X > b)$.

The above bounds are valid also when in (i) $\{X \leq b\}$ is replaced by $\{X < b\}$, and when in (ii) $\{X > b\}$ is replaced by $\{X \geq b\}$.

Proof. Assume that ϕ is positive and increasing. Denote $C = \{X \leq b\}$. Then the first claim follows after observing that

$$\begin{aligned} E\phi(X)1_{A \cap B} - E\phi(X)1_{A \cap C} &= E\phi(X)1_{A \cap B \cap C^c} - E\phi(X)1_{A \cap B^c \cap C} \\ &\geq \phi(b)P(A, B, C^c) - \phi(b)P(A, B^c, C) \\ &= \phi(b)[P(A, B) - P(A, C)]. \end{aligned}$$

The second claim follows by observing that when ϕ is positive and decreasing, the above inequality is valid for $C = \{X > b\}$. \square

Lemma B.3. *Let X be a random number with a continuous cumulative distribution function F . Then for any $a \in \mathbb{R}$, any $p \in [0, \infty)$, and any random number Y defined on the same probability space as X ,*

$$E|1(X > a) - 1(Y > a)| \leq c b(a)^{\frac{p}{1+p}} (E|X - Y|^p)^{\frac{1}{p+1}},$$

where

$$b(a) = \sup_{x: x \neq a} \frac{|F(x) - F(a)|}{|x - a|}.$$

and $c = 2^{p/(1+p)}(1+p)^{1/(1+p)}$. Moreover, the same upper bound is also valid for $E|1(X \geq a) - 1(Y \geq a)|$.

Proof. Observe that

$$E|1(X > a) - 1(Y > a)| = P(Y \leq a < X) + P(X \leq a < Y). \quad (\text{B.1})$$

We will derive upper bounds for the probabilities on the right, one by one.

Assume first that $P(Y \leq a < X) < 1$. Then because the cumulative distribution function of X is continuous, we may fix a number $b \in [a, \infty)$ such that $P(Y \leq a < X) = P(a < X \leq b)$. Then by applying Lemma B.2:(i) with $\phi(x) = (x - a)^p 1(x > a)$, $A = \{X > a\}$ and $B = \{Y \leq a\}$ we find that

$$\begin{aligned} E|X - Y|^p 1(Y \leq a < X) &\geq E(X - a)^p 1(Y \leq a < X) \\ &\geq E(X - a)^p 1(a < X \leq b) \\ &= E(X - a)^p 1_{(a, b]}(X). \end{aligned}$$

Let Q be a quantile function of X , and note [FS04, Section A.3] that

$$\begin{aligned} E(X - a)^p 1_{(a, b]}(X) &= \int_{F(a)}^{F(b)} (Q(u) - a)^p du \\ &= \int_{F(a)}^{F(b)} \left(\frac{Q(u) - a}{u - F(a)} \right)^p (u - F(a))^p du \\ &\geq c_a^p \int_{F(a)}^{F(b)} (u - F(a))^p du \\ &= c_a^p \frac{(F(b) - F(a))^{p+1}}{p+1} = c_a^p \frac{P(a < X \leq b)^{p+1}}{p+1} \end{aligned}$$

where

$$c_a = \inf_{u \in (F(a), 1)} \left(\frac{Q(u) - a}{u - F(a)} \right).$$

After recalling that $P(Y \leq a < X) = P(a < X \leq b)$, we may conclude that

$$E|X - Y|^p 1(Y \leq a < X) \geq c_a^p \frac{P(Y \leq a < X)^{p+1}}{p+1}. \quad (\text{B.2})$$

The above inequality is also true in the case with $P(Y \leq a < X) = 1$, in which case we choose $b = \infty$ and replace $F(b)$ by 1 in the above derivation.

To obtain an upper bound for the second term on the right side of (B.1), assume that $P(X \leq a < Y) < 1$. Then by the continuity of the cdf of X , we may fix a number $c \leq a$ such that $P(X \leq a < Y) = P(c < X \leq a)$. Then by applying Lemma B.2:(ii) with $\phi(x) = (a - x)^p 1(x \leq a)$, $A = \{X \leq a\}$ and $B = \{Y > a\}$ we find that

$$\begin{aligned} E|X - Y|^p 1(X \leq a < Y) &\geq E(a - X)^p 1(X \leq a < Y) \\ &\geq E(a - X)^p 1(c < X \leq a) \\ &= E(a - X)^p 1_{(c, a]}(X). \end{aligned}$$

A similar computation as above using the quantile function shows that

$$\begin{aligned} E(a - X)^p 1_{(c, a]}(X) &= \int_{F(c)}^{F(a)} (a - Q(u))^p du \\ &= \int_{F(c)}^{F(a)} \left(\frac{a - Q(u)}{F(a) - u} \right)^p (F(a) - u)^p du \\ &\geq \hat{c}_a^p \int_{F(c)}^{F(a)} (F(a) - u)^p du \\ &= \hat{c}_a^p \frac{(F(a) - F(c))^{p+1}}{p+1} = \hat{c}_a^p \frac{P(c < X \leq a)^{p+1}}{p+1} \end{aligned}$$

where

$$\hat{c}_a = \inf_{u \in (0, F(a))} \left(\frac{a - Q(u)}{F(a) - u} \right).$$

After recalling that $P(X \leq a < Y) = P(c < X \leq a)$, we may conclude that

$$E|X - Y|^p 1(X \leq a < Y) \geq \hat{c}_a^p \frac{P(X \leq a < Y)^{p+1}}{p+1}. \quad (\text{B.3})$$

Again, the above inequality is also true in the case with $P(X \leq a < Y) = 1$, in which case we choose $c = -\infty$ and replace $F(c)$ by 0 in the above derivation.

By combining (B.2)–(B.3), then applying Jensen's inequality $(\frac{x+y}{2})^{1+p} \leq \frac{x^{1+p} + y^{1+p}}{2}$, and then equation (B.1), we find that

$$\begin{aligned} E|X - Y|^p &\geq c_a^p \frac{P(Y \leq a < X)^{p+1}}{p+1} + \hat{c}_a^p \frac{P(X \leq a < Y)^{p+1}}{p+1} \\ &\geq 2^{-p} (c_a \wedge \hat{c}_a)^p (p+1)^{-1} (P(Y \leq a < X) + P(X \leq a < Y))^{p+1} \\ &= 2^{-p} (c_a \wedge \hat{c}_a)^p (p+1)^{-1} (E|1(X > a) - 1(Y > a)|)^{p+1}. \end{aligned}$$

To conclude the claim, it suffices [FS04, Section A.3] to note that

$$c_a \wedge \hat{c}_a \geq \inf_{u \in (0,1): u \neq F(a)} \frac{|a - Q(u)|}{|F(a) - u|} \geq \left(\sup_{x: x \neq a} \frac{|F(x) - F(a)|}{|x - a|} \right)^{-1}.$$

By replicating the same proof with obvious modifications, one can verify that the same upper bound valid also for $E|1(X \geq a) - 1(Y \geq a)|$. \square

Proof of Proposition B.1. Assume first that $\epsilon > 0$. Denote by μ the variation measure of f . Note first that $|f(x_2) - f(x_1)| \leq \mu(x_1, x_2]$ for $x_1 \leq x_2$ and $|f(x_2) - f(x_1)| \leq \mu(x_2, x_1]$ for $x_1 > x_2$. We may write this estimate as

$$|f(x_2) - f(x_1)| \leq \int_{\mathbb{R}} [1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y)] \mu(dy).$$

We will split the above integral into two parts by expressing the above estimate as

$$|f(x_2) - f(x_1)| \leq I_1(x_1, x_2) + I_2(x_1, x_2),$$

where

$$I_1(x_1, x_2) = \int_{(-\epsilon, \epsilon)} [1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y)] \mu(dy)$$

and

$$I_2(x_1, x_2) = \int_{\mathbb{R} \setminus (-\epsilon, \epsilon)} [1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y)] \mu(dy).$$

By the Lipschitz continuity assumption we know that f has a derivative f' at almost every point in $(-\epsilon, \epsilon)$ bounded by a constant $c_1(f)$. The variation measure of f satisfies $\mu(dy) = |f'(y)|dy$ on $(-\epsilon, \epsilon)$, so that

$$I_1(x_1, x_2) \leq c_1(f) \int_{-\infty}^{\infty} [1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y)] dy = c_1(f) |x_1 - x_2|,$$

and

$$EI_1(X_1, X_2)^p \leq c_1(f)^p E|X_1 - X_2|^p \leq c_1(f)^p (E|X_1 - X_2|^q)^{p/q}. \quad (\text{B.4})$$

For the second term, we see by Jensen's inequality that

$$\begin{aligned} I_2(x_1, x_2)^p &\leq \mu(\mathbb{R} \setminus (-\epsilon, \epsilon))^{p-1} \int_{\mathbb{R} \setminus [-\epsilon, \epsilon]} [1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y)]^p \mu(dy) \\ &\leq \mu(\mathbb{R})^{p-1} \int_{\mathbb{R} \setminus (-\epsilon, \epsilon)} [1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y)]^p \mu(dy) \\ &= V(f)^{p-1} \int_{\mathbb{R} \setminus (-\epsilon, \epsilon)} [1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y)] \mu(dy). \end{aligned}$$

Observe next that $1_{(x_1, x_2]}(y) + 1_{(x_2, x_1]}(y) = |1(x_2 \geq y) - 1(x_1 \geq y)|$. Hence

$$EI_2(X_1, X_2)^p \leq V(f)^{p-1} \int_{\mathbb{R} \setminus (-\epsilon, \epsilon)} E|1(X_2 \geq y) - 1(X_1 \geq y)| \mu(dy).$$

Now by Lemma B.3,

$$E|1(X_2 \geq y) - 1(X_1 \geq y)| \leq c_3 b(y)^{\frac{q}{1+q}} (E|X_1 - X_2|^q)^{\frac{1}{q+1}},$$

where

$$b(y) = \sup_{x: x \neq y} \frac{|F_1(x) - F_1(y)|}{|x - y|},$$

and $c_3 = 2^{q/(1+q)}(1+q)^{1/(1+q)}$, and F_1 is the cumulative distribution function of X_1 . We conclude that

$$EI_2(X_1, X_2)^p \leq c_3 V(f)^p \left(\sup_{|y| \geq \epsilon} b(y) \right)^{\frac{q}{q+1}} (E|X_1 - X_2|^q)^{\frac{1}{q+1}}.$$

Next, observe that

$$\frac{|F_1(x) - F_1(y)|}{|x - y|} \leq 2\epsilon^{-1}$$

for all y and all x such that $|x - y| \geq \epsilon/2$, and

$$\frac{|F_1(x) - F_1(y)|}{|x - y|} \leq \sup_{|x| > \epsilon/2} p_1(x) =: c_2(p_1)$$

for all $|y| \geq \epsilon$ and $|x - y| < \epsilon/2$. Hence $b(y) \leq 2\epsilon^{-1} + c_2(p_1)$ for all $|y| \geq \epsilon$. Therefore,

$$EI_2(X_1, X_2)^p \leq c_3 V(f)^p (2\epsilon^{-1} + c_2(p_1))^{\frac{q}{q+1}} (E|X_1 - X_2|^q)^{\frac{1}{q+1}}. \quad (\text{B.5})$$

Now by noting that

$$E|f(X_1) - f(X_2)|^p \leq 2^{p-1} \left\{ EI_1(X_1, X_2)^p + EI_2(X_1, X_2)^p \right\}$$

and collecting the terms in (B.4) and (B.5) into the constants c , $d_1(f)$, and $d_2(p_2)$ in the statement, we obtain the claim for $\epsilon > 0$.

When $\epsilon = 0$, we can adapt the above proof easily to obtain the claim also in this case. It suffices to note that $I_1(X_1, X_2) = 0$, and $b(y) \leq \sup_{x \in \mathbb{R}} p_1(x)$ for all $y \in \mathbb{R}$.

□

C Elementary probability theory

The following elementary result, arguably well known to many, is included here for the reader's convenience.

Lemma C.1. *Let A_1, A_2, \dots be events such that $P(\cup_k A_k) = 1$. Consider random variables X, X_1, X_2, \dots such that for all $k \geq 1$, $1_{A_k} X_n \xrightarrow{P} 1_{A_k} X$ as $n \rightarrow \infty$. Then $X_n \xrightarrow{P} X$.*

Proof. Recall that $X_n \xrightarrow{P} X$ if and only if every subsequence $\mathbb{N}_0 \subset \mathbb{N}$ has a further subsequence along which $X_n \rightarrow X$ almost surely [Kal02, Lemma 4.2]. To prove the claim, let \mathbb{N}_0 be an arbitrary subsequence of \mathbb{N} . For $k = 1$, we may choose a subsequence $\mathbb{N}_1 \subset \mathbb{N}_0$ such that $1_{A_1} X_n \rightarrow 1_{A_1} X$ almost surely as $n \rightarrow \infty$ along \mathbb{N}_1 . For $k = 2$, may choose a further subsequence $\mathbb{N}_2 \subset \mathbb{N}_1$ such that $1_{A_2} X_n \rightarrow 1_{A_2} X$ almost surely along \mathbb{N}_2 . Proceeding this way, we construct a nested collection $\mathbb{N}_k \subset \mathbb{N}_{k-1} \subset \cdots \subset \mathbb{N}_0$ such that $1_{A_k} X_n \rightarrow 1_{A_k} X$ almost surely as $n \rightarrow \infty$ along \mathbb{N}_k .

Denote the elements of \mathbb{N}_k by $n_{k,1} < n_{k,2} < \cdots$, and define $n_\ell = n_{\ell,\ell}$. Then for all $k \geq 1$, $n_\ell \in \mathbb{N}_k$ for all $\ell \geq k$, so that $1_{A_k} X_{n_\ell} \rightarrow 1_{A_k} X$ almost surely as $\ell \rightarrow \infty$. Hence we conclude that $1_A X_{n_\ell} \rightarrow 1_A X$ almost surely where $A = \cup_k A_k$. Because $P(A) = 1$, we conclude that $P(X_{n_\ell} \rightarrow X) = P(1_A X_{n_\ell} \rightarrow 1_A X) = 1$. Hence $\mathbb{N}'_0 = \{n_\ell : \ell \geq 1\}$ is a subsequence of \mathbb{N}_0 along which $X_n \rightarrow X$ almost surely. Hence $X_n \xrightarrow{P} X$. \square

D Auxiliary facts

The lemma helps shows that on the average, singularities do not matter too much, if X has a bounded probability density function. The following lemma has appeared in [Tik12, Lemma 3.2]. For the reader's convenience, we include a short proof of the result below.

Lemma D.1. (i) Let X be a random number with a bounded probability density function p_X . Then for any real number y and any $\alpha \in (0, 1)$,

$$E|X - y|^{-\alpha} \leq 1 + \frac{2}{1 - \alpha} \sup_x p_X(x).$$

(ii) Let (X_t) be a real-valued random process such that X_t has a bounded density p_{X_t} for every t . Then for any measurable function (y_t) ,

$$E \int_0^T |X_t - y_t|^{-\alpha} dt \leq T + \frac{2}{1 - \alpha} \int_0^T \sup_x p_{X_t}(x) dt.$$

Proof. (i) Denote $c = \sup_x p_X(x)$ and note that $E|X - y|^{-\alpha} 1_{\{|X - y| > 1\}} \leq 1$ and

$$E|X - y|^{-\alpha} 1_{\{|X - y| \leq 1\}} = \int_{-1}^1 |x|^{-\alpha} p_X(x + y) dx \leq c \int_{-1}^1 |x|^{-\alpha} dx = \frac{2c}{1 - \alpha}.$$

(ii) This follows at once from (i). Naturally, we assume in part (ii) that the density of X_t is measurable as a function of t . \square

References

- [AMV10] E. Azmoodeh, Y. Mishura, and E. Valkeila. On hedging European options in geometric fractional Brownian motion market model. *Statistics and Decisions*, 27:129–143, 2010.

- [ASVY14] Ehsan Azmoodeh, T. Sottinen, L. Viitasaari, and A. Yazigi. Necessary and sufficient conditions for Hölder continuity of Gaussian processes. *Statist. Probab. Lett.*, 94:230–235, 2014.
- [AV15] E. Azmoodeh and L. Viitasaari. Rate of convergence for discretization of integrals with respect to fractional Brownian motion. *J. Theor. Probab.*, 28(1):396–422, 2015.
- [Avi09] Rainer Avikainen. On irregular functionals of SDEs and the Euler scheme. *Finance and Stochastics*, 13(3):381–401, 2009.
- [BN11] M. Besalú and David Nualart. Estimates for the solution to stochastic differential equations driven by a fractional Brownian motion with Hurst parameter $H \in (\frac{1}{3}, \frac{1}{2})$. *Stochastics and Dynamics*, 11:243–263, 2011.
- [CKM03] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima. Fractional Ornstein–Uhlenbeck processes. *Electron. J. Probab.*, 8(3), 2003.
- [DN98] R. M. Dudley and R. Norvaiša. *An introduction to p-variation and Young integrals*. MaPhySto Lecture Notes, Vol I, 1998.
- [DNPV12] Eleonora Di Nezza, Giampiero Palatucci, and Enrico Valdinoci. Hitchhiker’s guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136:521–573, 2012.
- [FH14] Peter K. Friz and Martin Hairer. *A course on rough paths*. Springer, 2014.
- [FS04] Hans Föllmer and Alexander Schied. *Stochastic Finance*. Walter de Gruyter & Co., 2004.
- [Kal02] Olav Kallenberg. *Foundations of Modern Probability*. Springer, 2002.
- [KD92] Harold J. Kushner and Paul G. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, 1992.
- [Lyo98] T.J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.
- [NR02] D. Nualart and A. Răşcanu. Differential equations driven by fractional Brownian motion. *Collect. Math.*, 53:55–81, 2002.
- [Sam91] G. Samorodnitsky. Probability tails of Gaussian extrema. *Stochastic Processes and Their Applications*, 38(1):55–84, 1991.
- [SKM93] S.G. Samko, A.A. Kilbas, and O.I. Marichev. *Fractional Integrals and Derivatives: Theory and Applications*. Gordon and Breach Science Publishers, 1993.

- [Son06] D. Sondermann. *Introduction to Stochastic Calculus for Finance: A New Didactic Approach*. Springer, 2006.
- [Tik12] Heikki Tikanmäki. Integral representations of some functionals of fractional Brownian motion. *Commun. Stoch. Anal.*, 6(2):193–212, 2012.
- [You36] L.C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta. Math.*, 67:251–282, 1936.
- [Zäh98] M. Zähle. Integration with respect to fractal functions and stochastic calculus. I. *Probab. Theory Relat. Fields*, 111:333–372, 1998.
- [Zäh01] M. Zähle. Integration with respect to fractal functions and stochastic calculus. II. *Math. Nachr.*, 225:145–183, 2001.